Imaging In A Distorting Medium

A thesis
submitted for the degree of
Doctor of Philosophy in Electrical Engineering
from the
University of Canterbury,
Christchurch, New Zealand.

R. A. Minard, B.Sc.(Hons)

1985

ABSTRACT

Images of an object embedded in an inhomogeneous medium are shown to contain speckle artefacts. The efficacy of a speckle processing technique called 'shift-and-add' for removing these artefacts is investigated.

Shift-and-add is known to be applicable to any object which is dominated by a bright isolated point scatterer. This scatterer is called the 'reference'. Evidence that certain naturally occurring biological structures are capable of acting as a reference is presented. It is shown that an idealised version of shift-and-add, in conjunction with Fourier imaging, provides an approximate method for imaging an arbitrary object. Experimental results are reported which show that under most circumstances shift-and-add is able to reduce artefacts caused by multiple scattering.

When the object is not dominated by point scatterers the shift-and-add image contains a significant 'ghosting' artefact. A theoretical model which accounts for this artefact is proposed and verified by simulation. Several extensions of shift-and-add are reported which allow the artefacts to be reduced and so allow an appreciably larger class of objects to be imaged. In particular, most objects dominated by several bright point scatterers can be reconstructed. This is verified experimentally for a simple object embedded in animal tissue. Additionally, a wide class of piecewise-constant extended objects can be reconstructed. One dimensional simulations are presented to support this conclusion.

A generalised shift-and-add principle, which in theory allows arbitrary objects to be reconstructed, is proposed. It is shown that, in practice, the class of objects which can be reconstructed by the extensions to shift-and-add introduced in this thesis is considerably wider than that which can be successfully imaged by previously reported versions of shift-and-add.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am grateful to my supervisor Professor Richard Bates for his considerable inspiration, guidance and encouragement.

During the course of this work, I was supported by a Postgraduate Scholarship from the New Zealand University Grants Committee.

I thank my colleagues in the Electrical and Electronic Engineering Department for providing useful discussions and a pleasant enviroment to work in. In particular, I thank Alastair Sinton for his assistance with the preparation of the thesis.

PREFACE

During the past decade a revolution has occurred in the way in which measurements are made in the applied sciences. Advances in technology have so expanded measurement objectives that a new science, which might appropriately be called 'computer-aided imaging', has come into being.

The changes in objectives which have occurred are most apparent in the sophistication of the models that are now being invoked to characterise physical systems. In the days when measurements were made manually there was only enough data to justify simple models, typically a functional relationship with a few free parameters to be determined by a model fitting procedure. Thus a physical system was described by a law and the values of a few physical constants. It was felt that the functional form of the law was more significant than the values of the constants which were seen as largely incidental.

With the advent of computer-controlled data acquisition apparatus it has become possible to obtain orders of magnitude more data measurements and so to fit more parameters to the model. In a model determined by a large number of parameters the relationship of the system to the model is mainly governed by the values of the parameters. The outcome of this trend is that the importance of the 'law' has declined so that most of the information is conveyed by the parameters. The system is then represented at least in the first instance by an array of samples and a sampling function. When the model is considered as a whole, the form of the sampling function is not very important. Conceptually the measurements are now more readily interpreted as an image rather than as a list of numbers.

This thesis is concerned with computer-aided imaging in an inhomogeneous medium. Conventional techniques assume that the medium is homogeneous. Consequently, the images of an object embedded in an inhomogeneous medium obtained with such techniques are considerably distorted. One area in which this problem has been addressed is optical astronomy (Bates 1982). A class of techniques called 'speckle processing' has been developed to restore images of astronomical objects obtained through a turbulent (and therefore inhomogeneous) atmosphere. Recently, speckle processing techniques have been applied to ultrasonic imaging through mammalian tissues. In particular, a technique called 'shift-and-add' has been found to be effective in reducing the distortions. Preliminary results of this work are presented by Bates and Robinson (1981) and Robinson (1982). Faithful images of the object are obtained by

shift-and-add only when the object contains a strong point scatterer. This scatterer is called the 'reference'.

In this thesis, the work of Robinson (1982) is considerably extended. A firm theoretical basis for speckle processing in ultrasonic media is presented, and additional experimental results for more realistic distortions than those employed by Robinson (1982) are obtained. Extensions of the technique are formulated and shown to significantly improve the quality of the restored image. The extensions are expected to be applicable to optical astronomical speckle processing.

The thesis is divided into three parts. Part I introduces computer-aided imaging techniques. Part II presents new results, and Part III contains conclusions and suggestions for future research.

Part I comprises three chapters. Chapter 1 describes the algorithms employed by conventional computer-aided imaging systems. A mathematical basis is established for image processing. The physics of wave propagation in an inhomogeneous medium are reviewed. An object may be measured directly or by the radiation it scatters or emits. The two methods are compared. A beam of radiation may be used to probe the object. The nature of the object determines how it scatters the beam. The 'remote probing' problem consists of inferring the nature of the object from the observed scattering. The problem is defined in detail. Approximate methods for calculating the scattering when the nature of the object is known are discussed. A solution, called 'Fourier imaging', of the remote probing problem based on these approximate methods is presented. In most situations, the distortion caused by an inhomogeneous medium is manifested in the distorted image as a 'speckly' pattern. This phenomenon is called 'speckle'.

In Chapter 2 the current status of ultrasonic imaging is assessed. The method of echo-location, which is the basis of radar and sonar and is also employed by medical diagnostic imaging systems, is derived. When the waves used to probe the object travel along straight rays, the method called 'computed tomography' represents an excellent solution of the remote probing problem. Ultrasound, however, travels along curved ray paths and is affected by diffraction. The consequences of this for practical imaging systems are explained. The literature contains several accounts of attempts at compensating for ray curvature. These attempts are described and compared. The method of diffraction tomography accounts for first order diffraction effects but not for ray curvature. Iterative procedures have been devised for accounting for the latter effect but are not very

successful. Finally, the speckle processing techniques, on which this thesis is based, are discussed.

Chapter 3 discusses the development of software for computer-aided imaging systems. Since much of this software development is for research purposes, it is subject to considerable modification. Techniques for designing software to facilitate such modification are described. It is shown that a special purpose software tool, for testing such programs, can be of considerable assistance. If, in addition, the tool is integrated with the developed software, several benefits accrue, as is explained in detail.

Part II contains new results obtained by the author. Chapter 4 describes a software module, called the 'GEM', written by the author to meet the demands outlined in Chapter 3. The module was developed using the methodology expounded in Chapter 3. All of the software developed by the author to implement the algorithms reported in the thesis use the GEM.

An ultrasonic measurement facility, used by the author to obtain experimental data for the thesis, is described in Chapter 5. During the author's research the computer originally used to control the facility was scrapped. To obtain further experimental data, the facility was interfaced to the Electrical and Electronic Engineering Department's new computer which is a Digital Equipment Corporation VAX 11/750. The design and construction of the interface is described.

Chapter 6 presents new developments of the shift-and-add principle. Experimental data indicate that certain biological structures are able to act as a reference for shift-and-add. Shift-and-add is shown to be an approximate solution of the remote probing problem. Some of the deficiencies of shift-and-add are discussed with the aid of a simple model of the technique devised by the author. Each speckle image consists of many randomly located and scaled replicas of the object. Shift-and-add attempts to find the brightest replica and superpose it with its fellows in the other speckle images. However, there is a significant probability that the location of the brightest replica will sometimes be determined incorrectly. This probability is calculated from the simple model.

Chapter 7 presents an extension, called 'correlation' shift-and-add, of ordinary shift-and-add which remedies some of its deficiencies. Essentially, the method uses cross-correlation to locate a replica of the object in each speckle image. An iterative technique is employed. The extension is analysed algebraically. Simulated examples and reconstructions obtained from experimental data are presented.

Three more extensions of shift-and-add are investigated in Chapter 8. The first, called 'inverse' shift-and-add, employs deconvolution rather than cross-correlation to locate the replicas of the object. The second, called 'differential' shift-and-add, may be applied to either correlation or inverse shift-and-add. A significant proportion of piecewise constant extended objects can be reconstructed by this procedure. An extension is proposed which promises to account for curvature of the probing beam as it propagates through the inhomogeneous medium.

The correlation, inverse, and differential extensions of shift-and-add are shown to be special cases of a general transform technique in Chapter 9. Also, the extensions are applicable not just to shift-and-add but also to the related averaging techniques of 'speckle masking' (Milner 1979) and 'LWH' processing (Lynds et al. 1976).

The final chapter of the thesis, Chapter 10, summarises the salient results reported herein and it also contains several suggestions for future research. Experience with a technique for constructing extended ultrasonic objects is discussed. A method for accelerating the convergence of the iterative techniques employed for the extensions is proposed. It is suggested how the simple model of shift-and-add propounded in Chapter 6 may be used to refine the iterative schemes investigated in Chapters 7, 8, and 9. A design for an ultrasonic phantom with a variable refractive index is suggested. The chapter closes with some final remarks concerning the status of shift-and-add imaging.

The work reported in this thesis was performed during the period from March 1981 to February 1985. During the course of the work the following papers and presentations have been prepared:

Fright W.R. and Minard R.A. 1983. "Optical and Digital Signal Processing", IPENZ Trans. 10/EMCh, 65-69.

Bates R.H.T. and Minard R.A. 1983. "Some new approaches to inverse scattering", Proc. SPIE Conf. Inverse Optics, 413, 56-60.

Bates R.H.T. and Minard R.A. 1984. "Compensation for multiple reflection", IEEE Trans. SU-31, 330-336.

Minard R.A., Robinson B.S. and Bates R.H.T. 1985. "Full-wave computed tomography. Part 3: Coherent shift-and-add imaging", Proc. IEE 132, 50-58.

The following paper is to be submitted for the SPIE August 1985 International Conference on Speckle.

Bates R.H.T., Sinton A.M. and Minard R.A. 1985. "Generalisation of shift-and-add imaging".

GLOSSARY

Certain simple functions occur often in computer-aided imaging. They are:

$$\operatorname{sinc}(x) = \sin(\pi x)/(\pi x)$$

$$\operatorname{rect}(x) = \begin{cases} 1 & -\tfrac{1}{2} \leq x < \tfrac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$\operatorname{gau}(x) = \exp(-\pi x^2)$$

$$\operatorname{tri}(x) = \begin{cases} 1-|x| & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\operatorname{step}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The Dirac delta function $\delta(x)$ is also used.

PART I

REVIEW

1. INTRODUCTION

1.1 PREAMBLE

The problem of how to determine the nature of an object from measurements made some distance from it is of central importance in applied science. It may be called the 'remote measurement' problem. Computers are widely used in the solution of this problem and typically the solution is presented visually as an image. An appropriate term for the class of procedures employed in processing and displaying the results of remote measurement is therefore 'computer-aided imaging (CAI)'.

The remote measurement problem is referred to as the remote 'sensing' problem or the remote 'probing' problem depending on the source of the energy which is detected. The term remote sensing applies to situations in which the object to be imaged is actively emitting radiation and the radiation cannot be altered to facilitate the measurements. In remote probing, however, the object usually has no energy sources within it and must therefore be probed with some radiation. Part of the remote probing problem is the selection of the form of the probing radiation to give the best possible results. This thesis is concerned with the remote probing problem, but since the remote sensing problem can be considered to be a part of the remote probing problem, it has relevance to the remote measurement problem in general.

The purpose of this thesis is to describe and demonstrate the potential usefulness of a CAI method for solving the remote probing problem when the object is embedded in an inhomogeneous medium. In addition, a mathematical and physical basis is provided for the method, to indicate its range of validity. The relevant background theory is developed in this chapter.

In ¶1.2, the general mathematical concepts of spaces and mappings are introduced for defining a terminology for CAI. A physical basis for remote probing in an inhomogeneous medium is derived from the scalar wave equation in ¶1.3. Direct and remote measurements are compared in ¶1.4. A detailed definition of the remote probing problem and the format of a general CAI solution to it is provided in ¶1.5. When the nature of the object is known, the way in which the probing radiation is scattered by the object may be calculated by various methods. The more important of these are discussed in ¶1.6. In ¶1.7, one of the methods for calculating the

scattering is used to formulate a widely used solution of the remote probing problem, called 'Fourier imaging'. Algorithms for implementing Fourier imaging are derived in ¶1.8. The images obtained from a CAI system based on Fourier imaging are usually distorted by a phenomenon called 'speckle'. The nature of the distortion and its origin are discussed in ¶1.9. Finally, the review material in this chapter is summarised and is related to the remainder of the thesis in ¶1.10.

1.2 A MATHEMATICAL BASIS FOR IMAGING

1.2.1 Semantics

In this section an overview of the CAI process is given and some relevant terminology is defined. The goal of CAI is to make information about a particular object available in a comprehensible format to an input port of a system. The system is referred to as the 'recipient'. Some examples of ports and their associated recipients are listed in Table 1.1.

Table 1.1 Some examples of ports and their associated recipients.

Recipient	Port
A digital computer.	A memory location which can be altered independently of the computer's central processing unit (CPU).
Photographic film.	The emulsion of the film.
The human brain.	Any one of the senses.

The object can be completely described by the way in which its material properties are distributed in space and time. Generally, only a subset of these properties are considered in a given CAI situation and usually only in a limited volume and for a limited time. For instance, in ultrasonic imaging the electromagnetic properties such as the dielectric constant are ignored.

The first step in CAI is measurement. Fields which have been perturbed by the object, and are therefore carrying information about the object, are detected. The next step is the transformation of the field

measurements into values which represent the object's properties. This is termed 'image reconstruction'. These values are then converted into a form which can be received by the input port. This final step is referred to as 'image display'. This decomposition of CAI is based upon the concept of information flow. Information about the object is transformed in various steps until it can be accepted by the input port. Fig. 1.1 illustrates this.

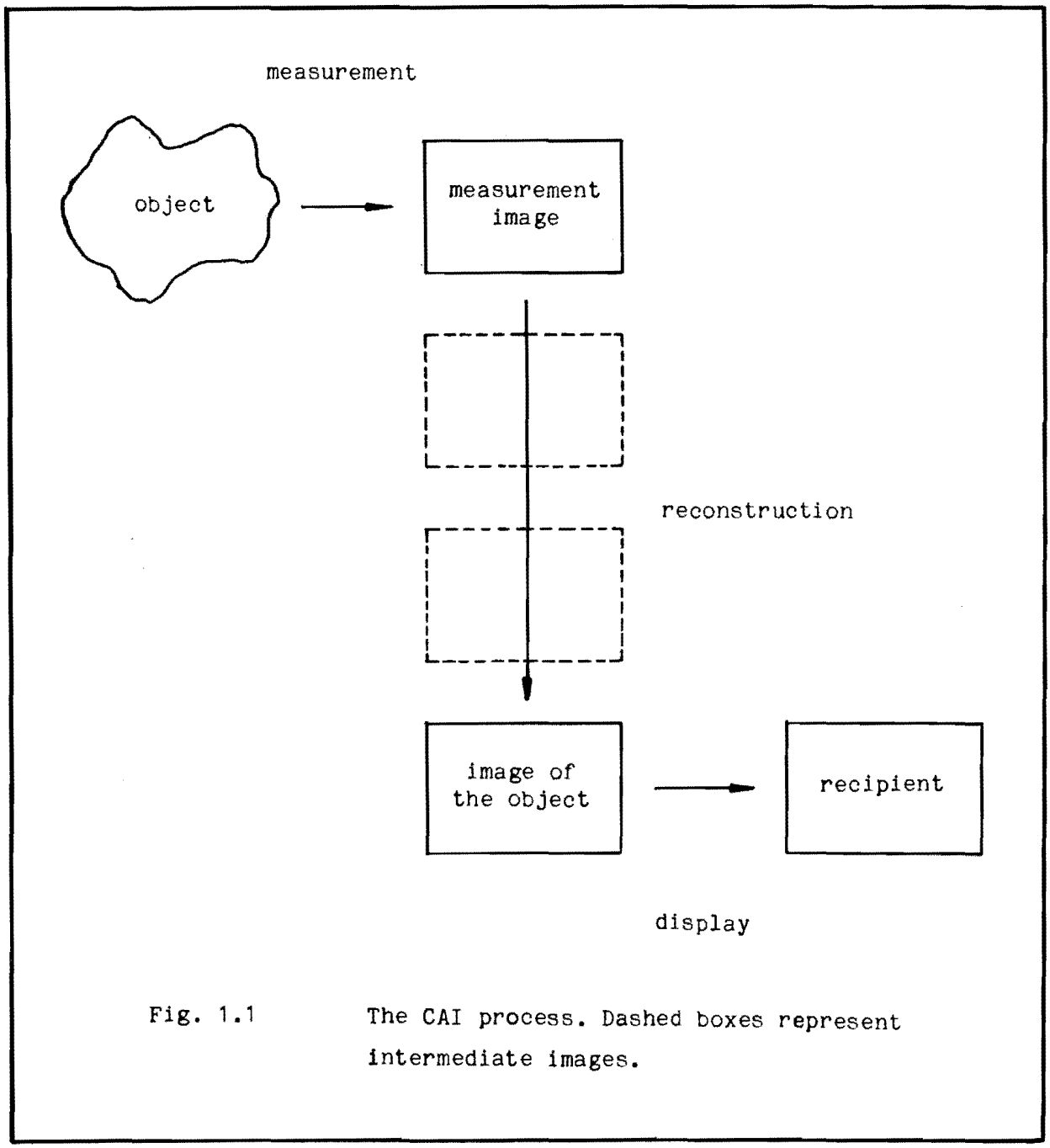


Fig. 1.1 The CAI process. Dashed boxes represent intermediate images.

Although the word 'image' has been used in the terminology defined in the previous paragraph, it has not been defined. Conventionally, an 'image' is defined to be:

'a visual representation of an object'. (1.1)

Photographs and video sequences fit this definition. However, there are representations from some measurement systems which do not. For instance, consider the use on board a ship of an echo-sounder to determine sea depth. The object to be imaged is the water beneath the ship and the property of interest is the distance to the sea floor. Someone listening to the sequence of reverberations from the echo-sounder is able to picture the variation in depth as the ship sails along. This is clearly an example of remote measurement yet no visual representation is present. If an image is considered to have been formed in this situation, definition (1.1) is inadequate. An appropriate alternative definition is that:

'information about the object constitutes an image when that information is within the recipient'. (1.2)

When the image reconstruction step is examined in more detail for a particular situation, it can frequently be divided up into steps, in which the information describing the object is represented by a function of several variables, called an 'intermediate' function. In a similar manner the measurements are functions of time and space, as is the result of image reconstruction. Additionally, the reconstructed image has the important characteristic that it is representative of the properties of the object which are to be displayed. Now the result of each step can be displayed in a similar manner to the reconstructed image. This is obvious when the function concerned is dependent only on spatial and temporal variables. When other variables such as frequency are present these are identified with one of the space and time coordinates for the purpose of display. In some cases the function has a physical interpretation and so is more than a mathematical abstraction. As an example, consider a CAI process based on ultrasonic wave scattering and using Fourier techniques (Wolf 1969). The measurements consist of the values of the scattered wave at many times and places. One intermediate step is the calculation of the far-field radiation pattern of the object. This can be visualised and can seem to be just as tangible as its optical counterpart, the diffraction pattern, which one can see directly. Each of these intermediate functions therefore has many properties in common with the reconstructed image. With this in mind, an 'image' is defined to be:

'any function occurring in the CAI process which represents the information about the object', (1.3)

and an 'image of the object' is defined to be:

'an image which is mathematically close to the object's properties'. (1.4)

Just what 'mathematically close' means is specified in ¶1.2.3.

1.2.2 Mathematical Formulation

With definitions (1.3) and (1.4) in mind, a mathematical formulation of CAI can be constructed. One way to build the description is to start by characterising the kinds of measurements which can be made. There are two fundamental aspects to a measurement. The first is where and when it is made, and the second is the result. The former aspect can be termed the 'measurement event'. The set of all events is denoted by T_X and is called 'physical space'. If T_X has N spatial dimensions, N+1 real coordinates are needed to identify a particular event X. One of these specifies the time t at which the measurement is made,

$$x_0 = c_0 t \ , \tag{1.5}$$

where c_0 is a constant velocity and the remainder, $x_1, x_2, \ldots, x_N$, specify the position. In actuality, the measurement must be made over a time interval and a spatial volume since otherwise no energy can be transferred to the detector. The velocity c_0 is chosen to suit the type of CAI being used. In optics it could be the speed of light in free space or in ultrasonics it might be the speed of sound in air or water. Vectors provide a convenient notation for the coordinates of an event. The event vector corresponding to X is,

$$\mathbf{x} = (x_0, x_1, \ldots, x_N) \ . \tag{1.6}$$

It is often convenient to denote an event by its event vector. The set of all possible event vectors is an (N+1)-dimensional vector space. If the real numbers are denoted by ¶R, this space is denoted by $¶R^{N+1}$. The dimension of any vector space A is written dim(A) and is the number of coordinates required to specify a vector uniquely within A. For example,

$$\dim(¶R^{N+1}) = N+1 \ . \tag{1.7}$$

A position vector **r** may be defined from the spatial coordinates of X to be

$$\mathbf{r} = (x_1, x_2, \ldots, x_N) \ . \tag{1.8}$$

The set of all positions is denoted by $T_{\mathbf{r}}$. For two event vectors $\mathbf{x}$ and $\mathbf{y}$ in T_X, the scalar product is defined by

$$\mathbf{x}\cdot\mathbf{y} = \sum_{i=0}^{N} x_i y_i \ . \tag{1.9}$$

The Euclidean distance between two events, X and Y, is

$$d(X,Y) = d(\mathbf{x},\mathbf{y}) = \left[\sum_{i=0}^{N} (x_i - y_i)^2 \right]^{1/2} , \tag{1.10}$$

where $\mathbf{y}$ is the event vector for Y. The function d is called the 'metric' for T_X and with d defined on it, T_X may be referred to as a 'metric space' (Burkill and Burkill 1970). Note that this metric is not the Lorentz metric (Morse and Feshbach 1953 ¶1.7) used in special relativity but has been constructed to allow any combination of spatial coordinates and the time coordinate to be used as the image coordinates.

In some situations it is desirable to redefine the coordinates for each point. For instance in an inhomogeneous medium the ray paths are curved and a coordinate system which follows the curvature may be more useful. A coordinate transformation e, between the two systems may be written as

$$\mathbf{x}' = e(\mathbf{x}) \ , \tag{1.11}$$

where $\mathbf{x}'$ is the new event vector for X. In general the form of the metric d is different in the transformed coordinate system. The new form is found using

$$d(X,Y) = d(\mathbf{x},\mathbf{y}) = d'(\mathbf{x}',\mathbf{y}') \ , \tag{1.12}$$

where, in the transformed system, d' is the metric and $\mathbf{y}'$ is the event vector for Y. A transformation for which the form of the metric is unchanged,

$$d'(\mathbf{x}',\mathbf{y}') = d(\mathbf{x}',\mathbf{y}') \ , \tag{1.13}$$

is called a 'coordinate isometry'. Examples of coordinate isometries are translations, rotations, reflections and time reversals.

The second aspect of a measurement, the result, can always be described by a certain set of real numbers. For scalar quantities such as the electric charge density, one number suffices. A vector quantity has N components. For instance, in three dimensions the electric field vector has

three components. A second rank tensor such as the elasticity tensor has 9 components in three-dimensional space. The set of all possible values for the result of a measurement may be called the 'amplitude' space Ω of the measurement. The amplitude spaces for the previous examples are $\mathbb{R}$, $\mathbb{R}^3$ and $\mathbb{R}^9$ respectively.

The mathematical definition of an image used in this thesis is:

'an image is any function defined on a subset of a metric space'. (1.14)

For a subset T_{sub} of T_X, and an image f, this is written

$$f : T_{sub} \rightarrow \Omega \tag{1.15}$$

using mapping notation. The colon and arrow in (1.15) indicate that each event in T_{sub} is mapped by f to an amplitude in Ω. An important aspect of this definition is that T_{sub} belongs to a metric space. This allows spatial and temporal relationships to be quantified, both within an image and between an image and an object. The set of all possible images with coordinate system T_{sub} and amplitude space Ω is written $\Lambda(T_{sub},\Omega)$. The set of all possible images for any coordinate system and any amplitude space is denoted by $\mathbf{\Lambda}$.

Images are classified by the nature of their coordinate and amplitude spaces. When the space T_X has a high dimension there are many possibilities for the dimensionality of T_{sub}. Three cases occur regularly enough to be worth defining explicitly, however. They are volume, surface and line images. The dimensions for these cases are

$$\dim(T_{sub}) = \dim(T_X) , \tag{1.16}$$

$$\dim(T_{sub}) = \dim(T_X) - 1 , \tag{1.17}$$

and

$$\dim(T_{sub}) = 1 \tag{1.18}$$

respectively. When the time coordinate is neglected these definitions are conventional in the sense that in three dimensions a volume image is three-dimensional and a surface image is two-dimensional.

Surface and line images are frequently used for image display purposes. An optical volume image could be formed from an infinite number of optical surface images placed in a pile. If the surface images are opaque to light it is only possible to observe the top one. Thus, it can be seen that surface images are the highest dimensionality images that can be

viewed on an opaque image display. Line images can be depicted on two-dimensional displays as a graph of image amplitude versus position.

Images are further classified on the basis of their amplitude spaces in the same way that measurements are. When the dimension of Ω is 0, 1 and 2, $\Lambda(T_{sub},\Omega)$ consists of scalar, vector and second rank tensor images respectively. Each component of the amplitude belongs to a set of numbers, e.g. the reals, the complex numbers or the integers. The name of this set is called the 'type' of the image. For example, an integral image has amplitude components which are integers. Such an image might arise in a photon-counting experiment.

Complex images can sometimes be used to simplify the processing of quantities which vary with time. For instance, when a scalar field is measured, the result can be thought of as an image $\psi(\mathbf{r},t)$ of the field in the measurement volume and over the measurement period. When the field can be written

$$\psi(\mathbf{r},t) = A(\mathbf{r}) \cos[\omega t + \phi(\mathbf{r})] , \tag{1.19}$$

it is referred to as a 'monochromatic' field of angular frequency ω, amplitude $A(\mathbf{r})$ and phase $\phi(\mathbf{r})$. The complex image $\psi(\mathbf{r})$ defined by

$$\psi(\mathbf{r}) = A(\mathbf{r}) \exp[j\phi(\mathbf{r})] \tag{1.20}$$

represents $\psi(\mathbf{r},t)$ since $\psi(\mathbf{r},t)$ is given in terms of $\psi(\mathbf{r})$ by

$$\psi(\mathbf{r},t) = \mathrm{re}[\psi(\mathbf{r}) \exp(j\omega t)] \tag{1.21}$$

where re(•) denotes the real part (the imaginary part is denoted by im(•) and the phase by phase(•)). A real monochromatic image $\psi(\mathbf{r},t)$ may therefore be represented by a time-independent complex image $\psi(\mathbf{r})$.

1.2.3 Image Operations

'Image processing' is the application of a sequence of image operations. An 'image operation' acts on a set of images to produce another. For instance, an operation B is called 'binary' when it acts on two images to produce a third (Larsen 1969 ¶1.5). It is described symbolically by

$$B : \Lambda^2 \to \Lambda . \tag{1.22}$$

Elements of Λ^2 are called 'pairs' of images.

Operations which act on amplitude or coordinate space may be extended to act on images in a fairly obvious manner. As an example consider an 'amplitude' operation a, which maps one amplitude space Ω to

another Ω',

$$a : \Omega \rightarrow \Omega' \ . \tag{1.23}$$

An 'image amplitude' operation A can be defined which maps $\Lambda(T_{sub},\Omega)$ to $\Lambda(T_{sub},\Omega')$:

$$A : \Lambda(T_{sub},\Omega) \rightarrow \Lambda(T_{sub},\Omega') \ . \tag{1.24}$$

For an image f in $\Lambda(T_{sub},\Omega)$, A(f) is an image in $\Lambda(T_{sub},\Omega')$ which is defined by

$$A(f) : T_{sub} \rightarrow \Omega' \tag{1.25}$$

and

$$A(f)(\mathbf{x}) = a\big(f(\mathbf{x})\big) \ . \tag{1.26}$$

A particular example of an amplitude operation on a complex image which may be extended in this way is complex conjugation.

Coordinate transformations may be extended in two ways, actively or passively (Bradley and Cracknell 1972 ¶1.4). An active transformation 'moves' the image while a passive one 'moves' the coordinate system. Given a coordinate transformation e on T_{sub},

$$e : T_{sub} \rightarrow T_{sub} \ , \tag{1.27}$$

an 'image coordinate transformation' E on $\Lambda(T_{sub},\Omega)$,

$$E : \Lambda(T_{sub},\Omega) \rightarrow \Lambda(T_{sub},\Omega) \ , \tag{1.28}$$

may be defined either 'actively' by

$$\big(E(f)\big)\big(e(\mathbf{x})\big) = f(\mathbf{x}) \tag{1.29}$$

or 'passively' by

$$\big(E(f)\big)\,(\mathbf{x}) = f\big(e(\mathbf{x})\big) \tag{1.30}$$

where f is an image in $\Lambda(T_{sub},\Omega)$ and e(f) is the result of applying E to f. The active transformation is usually more easily visualised and is adopted in this thesis. A coordinate isometry can be extended to an image isometry by using (1.27) to (1.30). An image isometry has the property of preserving the size of an image. For a class of physical systems known as 'point-spread invariant' (PSI) systems, certain local properties are the same from place to place. In such systems, which are defined in greater detail in ¶1.5.4, the 'translation image isometry' is obviously very important. The active image translation by a vector $\mathbf{y}$ is denoted by $T_{\mathbf{y}}$. The result of applying it to an image f is an image $T_{\mathbf{y}}(f)$ given by

$$T_{\mathbf{y}}(f)\,(\mathbf{x}) = f(\mathbf{x} - \mathbf{y}) \ . \tag{1.31}$$

Another important image operation is that of scaling. A scaling operator $S_{a_0,a_1,\ldots,a_N}$ is defined by

$$S_{a_0,a_1,\ldots,a} (f)\ (\mathbf{x}) = f\left(\frac{x_0,x_1,\ldots,x_N}{a_0,a_1,\ldots,a_N}\right) . \tag{1.32}$$

The images (of which f, g and h are typical examples) considered in the remainder of this section are taken to be complex. Two binary image operations that are useful when dealing with PSI systems are convolution and cross-correlation. Following Bracewell (1978), the convolution of f and g is defined by

$$(f \odot g)\ (\mathbf{x}) = \int_{T_X} f(\mathbf{x}')\ g(\mathbf{x}' - \mathbf{x})\ d\mathbf{x}' \tag{1.33}$$

and the cross-correlation by

$$(f * g)\ (\mathbf{x}) = \int_{T_X} f^*(\mathbf{x}' - \mathbf{x})\ g(\mathbf{x}')\ d\mathbf{x}' \tag{1.34}$$

where $d\mathbf{x}' = dx_0 dx_1 \ldots dx_N$. The image f * f is called the 'autocorrelation' of f. Some of the important properties of these operations are

$$f \odot g = g \odot f \tag{1.35}$$

$$(f \odot g) \odot h = f \odot (g \odot h) \tag{1.36}$$

$$f \odot (g + h) = f \odot g + f \odot h \tag{1.37}$$

$$(f \odot g)^\dagger = f^\dagger \odot g^\dagger \tag{1.38}$$

$$f * g = (g * f)^\dagger = f^\dagger \odot g \tag{1.39}$$

$$(f * g) * h = f \odot (g * h) \tag{1.40}$$

$$f * (g \odot h) = g \odot (f * h) \tag{1.41}$$

$$f * (g * h) = (f \odot g) * h = g * (f * h) \tag{1.42}$$

$$f * (g + h) = f * g + f * h \tag{1.43}$$

$$(f + g) * h = f * h + g * h \tag{1.44}$$

$$T_{\mathbf{x}}(f \odot g) = T_{\mathbf{x}}(f) \odot g = f \odot T_{\mathbf{x}}(g) \tag{1.45}$$

$$T_{\mathbf{x}}(f * g) = T_{-\mathbf{x}}(f) * g = f * T_{\mathbf{x}}(g) \tag{1.46}$$

$$\partial/\partial x_i (f \odot g) = \partial f/\partial x_i \odot g = f \odot \partial g/\partial x_i \tag{1.47}$$

$$\partial/\partial x_i (f * g) = -\partial f/\partial x_i * g = f * \partial g/\partial x_i \tag{1.48}$$

$$\int (f \odot g)\, dx_i = \int f\, dx_i \odot g = f \odot \int g\, dx_i \tag{1.49}$$

$$\int (f * g)\, dx_i = -\int f\, dx_i * g = f * \int g\, dx_i \tag{1.50}$$

where $\dagger$ is the adjoint operator (Herstein 1975 ¶6.8) defined by

$$f^{\dagger}(\mathbf{x}) = f^{*}(-\mathbf{x}) \tag{1.51}$$

and $\partial/\partial x_i$ denotes the operation of taking the partial derivative with respect to the i^{th} coordinate. The adjoint operator has the properties

$$(f^{\dagger})^{\dagger} = f\ , \tag{1.52}$$

$$T_{\mathbf{x}}(f^{\dagger}) = [T_{-\mathbf{x}}(f)]^{\dagger}\ , \tag{1.53}$$

$$\partial/\partial x_i(f^{\dagger}) = -(\partial f/\partial x_i)^{\dagger}\ , \tag{1.54}$$

and

$$\int f^{\dagger}\, dx_i = -\left(\int f\, dx_i\right)^{\dagger}\ . \tag{1.55}$$

Convolutions and cross-correlations can be converted to products by means of the 'Fourier' transformation. The Fourier transform of f, denoted by ¶F(f), is defined by

$$¶F(f)\ (\mathbf{u}) = \int_{T_{sub}} f(\mathbf{x})\ \exp(-j2\pi\mathbf{u}\cdot\mathbf{x})\ d\mathbf{x} \tag{1.56}$$

and its inverse may be deduced from the property

$$¶F(¶F(f)) = f^{\dagger *}\ . \tag{1.57}$$

It is

$$¶F^{-1}(f)\ (\mathbf{x}) = \int_{T_x} f(\mathbf{u})\ \exp(j2\pi\mathbf{u}\cdot\mathbf{x})\ d\mathbf{u}\ . \tag{1.58}$$

The Fourier transform relationship may also be written

$$f \leftrightarrow F$$

where the Fourier transform of a function is indicated by capitalising the letter representing the function. Using this notation some relationships involving the Fourier transform may be written as (Bracewell 1978 pp. 101-104,112)

$$f(\mathbf{x} - \mathbf{y}) \leftrightarrow \exp(-j2\pi\mathbf{u}\cdot\mathbf{y})\ F(\mathbf{u}) \tag{1.59}$$

$$f(a\mathbf{x}) \leftrightarrow F(\mathbf{u}/a)/|a| \tag{1.60}$$

$$R(f) \leftrightarrow R(F) \tag{1.61}$$

$$f^{\dagger} \leftrightarrow F^{*} \tag{1.62}$$

$$f \odot g \leftrightarrow FG \tag{1.63}$$

$$f * g \leftrightarrow F^{*}G \tag{1.64}$$

where a is a real constant, R is a rotation image isometry and $|\cdot|$ denotes the absolute value.

A 'distance' between two images, f and g, can be defined in a similar manner to the distance between two events. One such definition of a distance D(f,g) is

$$D(f,g) = \left[\int_{T_{sub}} |f(\mathbf{x}) - g(\mathbf{x})|^2 \, d\mathbf{x}\right]^{1/2} . \tag{1.65}$$

With this definition the distance between two images is zero if and only if the images are identical. In general, the smaller the distance between two images, the more closely they resemble each other.

1.2.4 Image Representations

When image manipulation is performed with the aid of a digital computer the image must be represented within a finite amount of storage. There are a wide variety of ways to do this, ranging from symbolic methods to sampling. In the 'symbolic method' the image is represented by a formula. Computer programs exist for the algebraic manipulation of such formulas (Pavelle et al. 1981; Hawkes 1983a, 1983b). As yet these have received little attention in image processing. A computer program written by the author which uses the symbolic representation is described in ¶7.4.1. The other extreme is the 'sampled representation' in which the image is represented by an array of numbers. When an image is input to a computer in analogue form the output of the sampling process is in just this form. This is one of the reasons why digital image processing uses the sampled representation almost exclusively.

'Image analysis' is the process of extracting information from an image. Often the process of image analysis can be thought of as the conversion of a sampled representation to a symbolic one. Consider, for instance, an image containing several peaks. When the amplitude and positions of the peaks have been found a simple formula can be written down to describe the image. The information originally represented by an array of samples is now represented by a formula.

1.2.5 Sampling Theory

A sampled representation of a function can be based on any property of the function. The conventional method which uses the value of the function is termed ordinate sampling. Some other properties which could be sampled are the slope or the extrema of a function, as indicated by Bracewell (1978).

When an image f is described by ordinate sampling it is represented by M samples $\{f_i\}$ obtained at events $\{\mathbf{x}_i\}$ where

$$f_i = f(\mathbf{x}_i) , \quad i = 1, \ldots, M . \tag{1.66}$$

The pair $(\mathbf{x}_i, f_i)$ is called a 'pixel'. A sampling function $s_i(\mathbf{x})$ must be defined for each pixel to complete the description. The image is reconstructed from its samples using the formula

$$f(\mathbf{x}) = \sum_{i=1}^{M} f_i \, s_i(\mathbf{x}) \quad . \tag{1.67}$$

For equation (1.67) to hold the sampling functions must have the property

$$s_i(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} = \mathbf{x}_i \\ 0 & \mathbf{x} = \mathbf{x}_j, \ 1 \leq j \leq M, \ j \neq i \end{cases} \tag{1.68}$$

for i = 1, ..., M. The form of the sampling functions can be simplified by employing sample events located on a rectangular grid. In this case the sampling functions are translated versions of a single prototype sampling function $s(\mathbf{x})$,

$$s_i(\mathbf{x}) = s(\mathbf{x} - \mathbf{x}_i) . \tag{1.69}$$

The position of the i^{th} sample $\mathbf{x}_i$ on the grid is specified by N + 1 indices $n_{i,j}$, j = 0, ..., N,

$$\mathbf{x}_i = (n_{i,0}\Delta_0, \ n_{i,1}\Delta_1, \ \ldots, \ n_{i,N}\Delta_N) \tag{1.70}$$

where Δ_j is the grid spacing for the j^{th} coordinate. The reconstruction formula (1.67) can be expressed as a convolution for regular sampling,

$$f(\mathbf{x}) = \Big[\sum_{i=1}^{M} f_i \delta(\mathbf{x} - \mathbf{x}_i) \Big] \odot s(\mathbf{x}) . \tag{1.71}$$

In two dimensions , with the time coordinate neglected, the sampling events $\mathbf{x}_i$ are described by sampling positions $\mathbf{r}_i$. The subscript i may be replaced by the subscript $(n_{i,1}, n_{i,2})$. The sampling positions $\mathbf{r}_i$ are then denoted by

$r_{n_{i,1},n_{i,2}}$, and the samples f_i of f by $f_{n_{i,1},n_{i,2}}$. The reference to the $n_{i,j}$ is no longer required, and f may be written as

$$f(\mathbf{x}) = \left[\sum_{j=1}^{J} \sum_{i=1}^{I} f_{i,j}\, \delta(\mathbf{x} - \mathbf{x}_{i,j}) \right] \circledcirc s(\mathbf{x}) \ . \qquad (1.72)$$

The samples $f_{i,j}$ are called an 'I × J' sampled representation of f. Exact sampled representations of f are possible under various conditions, the most important of which is when the Fourier transform of f is of limited extent (Bracewell 1978 p.194), i.e. when

$$\mathcal{F}(f)\ (\mathbf{u}) = 0 \qquad \text{for} \quad |u_i| \geq v_i/2 \ , \ i = 0, \ldots, N \ , \qquad (1.73)$$

for which the sampling functions are

$$s(\mathbf{x}) = \prod_{i=0}^{N} \operatorname{sinc}(v_i x_i) \ . \qquad (1.74)$$

When a function f satisfies (1.73) it is said to be 'band-limited'.

1.3 A PHYSICAL BASIS FOR ACOUSTICAL IMAGING IN AN INHOMOGENEOUS MEDIUM

In this section a wave equation is derived to describe the adiabatic propagation of longitudinal waves in a three-dimensional linear isotropic viscoelastic medium. The derivation begins in ¶1.3.1 by stating a constitutive equation which describes the acoustical properties of the medium. In ¶1.3.2 a vector wave equation is obtained by considering the constitutive equation together with momentum conservation. Longitudinal solutions of the wave equation are discussed in ¶1.3.3. In ¶1.3.4 the special case of an inviscid fluid is considered. The differences between wave propagation in an inhomogeneous medium and a homogeneous one are discussed in ¶1.3.5. Finally, integral representations of the field are described in ¶1.3.6.

1.3.1 The Constitutive Equation

The material properties of a physical system are described by a constitutive equation. The equation relates the state of the system to the forces which arise from that state. For a simple spring the constitutive equation relates the extension of the spring to the tension.

One method of specifying the state of a continuous medium is based upon considering it to be composed of many infinitesimally small particles. Each particle is referred to by specifying its position $\mathbf{r}_0$ at a time

t_0. There are two alternative ways of expressing a property f of the medium. In the Lagrangian description the property is expressed as a function of time for each particle, $f(\mathbf{r}_0,t)$, while in the Eulerian description the property is expressed as a function of position and time, $f(\mathbf{r},t)$.

The rate of change of the property can be calculated for a particular particle, df/dt, or at a fixed position, $\partial f/\partial t$. The two time derivatives are related by the equation,

$$df(\mathbf{r}_0,t)/dt = \partial f(\mathbf{r},t)/\partial t + \mathbf{v}(\mathbf{r},t)\cdot\nabla f(\mathbf{r},t) \ , \tag{1.75}$$

where $\mathbf{v}(\mathbf{r},t)$ is the velocity of particles at position $\mathbf{r}$, time t. As an illustration of the difference between the two derivatives consider the rate of change of position. At a fixed position,

$$\partial \mathbf{r}/\partial t = \mathbf{0} \tag{1.76}$$

while for a specified particle,

$$d\mathbf{r}/dt = \mathbf{v} \tag{1.77}$$

which is in accordance with (1.75).

The displacement of a particle from its reference position is called the 'strain', $\mathbf{s}(\mathbf{r}_0,t)$,

$$\mathbf{s}(\mathbf{r}_0,t) = \mathbf{r}(\mathbf{r}_0,t) - \mathbf{r}_0 \ . \tag{1.78}$$

Small strains are characterised by a strain tensor $\underline{\mathbf{S}}$ (Morse and Feshbach 1953 ¶1.6),

$$\underline{\mathbf{S}} = \left(\nabla\mathbf{s} + (\nabla\mathbf{s})^T\right)/2 \tag{1.79}$$

where the $(i,j)^{th}$ element of $\nabla\mathbf{s}$ is

$$(\nabla\mathbf{s})_{i,j} = \partial s_i/\partial x_j \tag{1.80}$$

and the superscript T denotes the transpose operation. A rate of strain tensor $\underline{\mathbf{V}}$ is constructed from the velocity in a similar manner,

$$\underline{\mathbf{V}} = \left(\nabla\mathbf{v} + (\nabla\mathbf{v})^T\right)/2 \tag{1.81}$$

with the property

$$\underline{\mathbf{V}} = d\underline{S}/dt \ . \tag{1.82}$$

In an ideal elastic medium a constant stress gives rise to a constant strain. This is in contrast to an ideal viscous medium, in which a constant stress gives rise to time-varying strain. However, the rate of

strain is constant. These observations lead to the formulation of the constitutive equation for a linear viscoelastic medium (Auld 1973 ¶3.E),

$$\underline{\mathbf{T}} = (\lambda|\underline{\mathbf{S}}| + \gamma|\underline{\mathbf{V}}|)\,\underline{\mathbf{I}} + 2\mu\underline{\mathbf{S}} + 2\eta\underline{\mathbf{V}} + \underline{\mathbf{F}} \tag{1.83}$$

where $\underline{\mathbf{T}}$ is the stress tensor, λ and γ are the Lamé constants of elasticity, μ, η are viscosity coefficients, $\underline{\mathbf{I}}$ is the unit tensor, $\underline{\mathbf{F}}$ represents any other stresses that may be present due to electromagnetic or gravitational effects and the trace of a tensor $\underline{\mathbf{A}}$ is denoted by $|\underline{\mathbf{A}}|$.

When the strain is time harmonic,

$$\mathbf{s}(\mathbf{r}_0,t) = \mathbf{s}(\mathbf{r}_0)\exp(-j\omega t) \tag{1.84}$$

where ω is the angular frequency, the velocity and the rate of strain tensor may be expressed in terms of the strain and the strain tensor:

$$\mathbf{v} = -j\omega\mathbf{s} \tag{1.85}$$

and

$$\underline{\mathbf{V}} = -j\omega\underline{\mathbf{S}}\ . \tag{1.86}$$

On substituting (1.85) and (1.86) into (1.83) the constitutive equation takes upon the simpler form,

$$\underline{\mathbf{T}} = \lambda'|\underline{\mathbf{S}}|\underline{\mathbf{I}} + 2\mu'\underline{\mathbf{S}} + \underline{\mathbf{F}}\ , \tag{1.87}$$

where

$$\lambda' = \lambda - j\omega\gamma \tag{1.88}$$

and

$$\mu' = \mu - j\omega\eta\ . \tag{1.89}$$

A viscoelastic medium subject to time harmonic strain can therefore be thought of as an elastic medium in which the Lamé constants are complex. The primes may be dropped without confusion since $\mu' = \mu$ and $\lambda' = \lambda$ for an ideal elastic medium.

1.3.2 The Wave Equation

The stress force per unit volume on the medium is the vector $\nabla\cdot\underline{\mathbf{T}}$, the i^{th} element of which is

$$(\nabla\cdot\underline{\mathbf{T}})_i = \sum_{j=1}^{3} \partial T_{ij}/\partial x_j\ . \tag{1.90}$$

The momentum per unit volume is $\rho\mathbf{v}$, where ρ is the density of the medium. The equation of motion relating the two quantities is

$$\nabla\cdot\underline{\mathbf{T}} = d(\rho\mathbf{v})/dt \ . \tag{1.91}$$

The requirement that the mass of the medium be conserved provides the equation of continuity,

$$\partial\rho/\partial t = -\rho\nabla\cdot\mathbf{v} \ . \tag{1.92}$$

Using (1.92) and (1.75), the rate of change of density is

$$d\rho/dt = -\rho\nabla\cdot\mathbf{v} + \mathbf{v}\cdot\nabla\rho \tag{1.93}$$

and the rate of change of momentum is

$$d(\rho\mathbf{v})/dt = -\rho\mathbf{v}\nabla\cdot\mathbf{v} + \mathbf{v}(\mathbf{v}\cdot\nabla\rho) + \rho d\mathbf{v}/dt \tag{1.94}$$

$$= \rho d\mathbf{v}/dt \ , \tag{1.95}$$

when terms of second order in the velocity are neglected. Similarly

$$d\mathbf{v}/dt = \partial\mathbf{v}/\partial t \tag{1.96}$$

to first order. Substituting (1.87) into (1.91), the time harmonic equation of motion is

$$\nabla(\lambda\nabla\cdot\mathbf{s}) + \nabla\cdot\left[\mu\left(\nabla\mathbf{s} + (\nabla\mathbf{s})^{T}\right)\right] + \nabla\cdot\underline{\mathbf{F}} = -\omega^2\rho\mathbf{s} \tag{1.97}$$

using the property

$$|\underline{\mathbf{S}}| = \nabla\cdot\mathbf{s} \ . \tag{1.98}$$

1.3.3 Longitudinal Waves

The vector wave equation (1.97) has both longitudinal and transverse solutions. When the shear modulus μ and coefficient of viscosity η are negligible, the transverse waves may be neglected. In this section only longitudinal or compression waves are considered.

For a longitudinal wave the strain field can be written as the gradient of a potential ϕ,

$$\mathbf{s} = -\nabla\phi \ . \tag{1.99}$$

Taking the divergence of (1.97) and substituting (1.99) into the result yields

$$\nabla^2(\lambda\nabla^2\phi) + \nabla\cdot\nabla\cdot(2\mu\nabla\nabla\phi) + \nabla\cdot\nabla\cdot\underline{\mathbf{F}} = -\omega^2\nabla\cdot(\rho\nabla\phi) \ . \tag{1.100}$$

After expanding the differential operators, a scalar wave equation

$$\begin{aligned}(\nabla^2 + \omega^2/c^2)\nabla^2\phi = -\big[&\nabla^2\lambda\nabla^2\phi + (\nabla\nabla 2\mu)\cdot(\nabla\nabla\phi) \\ &+ 2\nabla(\lambda + 2\mu)\cdot\nabla\nabla^2\phi \\ &+ \omega^2\nabla\rho\cdot\nabla\phi + \nabla\cdot\nabla\cdot\underline{\mathbf{F}}\big]/(\lambda + 2\mu)\end{aligned} \tag{1.101}$$

is obtained where the wave velocity c is given by

$$c^2 = (\lambda + 2\mu)/\rho \ . \tag{1.102}$$

The scalar product of two tensors, such as the second term inside the square brackets on the right hand side of equation (1.101), is defined by

$$\underline{\mathbf{A}}\cdot\underline{\mathbf{B}} = \sum_{j=1}^{N} \sum_{i=1}^{N} A_{ij}B_{ij} \ , \tag{1.103}$$

where $\underline{\mathbf{A}}$ and $\underline{B}$ are two second rank tensors.

Wave equations such as (1.101) are classified according to whether they describe propagation in a homogeneous or an inhomogeneous medium and hence whether the right-hand side is zero or non-zero respectively. The homogeneous version of (1.101) is

$$(\nabla^2 + \omega^2/c^2)\nabla^2\phi = 0 \ . \tag{1.104}$$

The wave propagation described by (1.101) can be thought of as that of (1.104) modified by the terms on the right-hand side of (1.101). These terms are therefore called 'source' terms since they give rise to the modified waves. The term $\nabla\cdot\nabla\cdot\underline{\mathbf{F}}$ is independent of the field and is therefore called a 'free' source term. The other, field dependent, terms are called 'induced' sources.

1.3.4 Longitudinal Waves in Inviscid Fluids

For an ideal inviscid fluid,

$$\mu = 0 \tag{1.105}$$

and

$$\lambda = 1/\kappa \tag{1.106}$$

where κ is the compressibility. The pressure p, which is given by the isotropic component of the stress tensor when there are no external forces, is found using (1.87), (1.99), (1.105), and (1.106) to be

$$p = (\nabla^2\phi)/\kappa \ . \tag{1.107}$$

The strain can be expressed in terms of the pressure by using equation (1.97). It is

$$\mathbf{s} = -\nabla\phi = (\nabla p - \nabla\cdot\underline{\mathbf{F}})/(\rho\omega^2) \ . \tag{1.108}$$

When (1.105) and (1.106) are substituted into (1.102) the acoustic velocity is found to be given by

$$c^2 = 1/(\rho\kappa) \ . \tag{1.109}$$

Upon substituting (1.107) and (1.108) into (1.101) and performing some manipulation, the wave equation for the pressure field is found to be

$$(\nabla^2 + \omega^2/c^2)p = (\nabla\rho\nabla p)/\rho + \left[\nabla\cdot(\rho\nabla\cdot\underline{\mathbf{F}})\right]/\rho \qquad (1.110)$$

which may be derived more directly from first principles but at the cost of some loss of insight. For instance, when a viscoelastic medium is assumed to be an inviscid fluid, (1.101) and (1.102) indicate the effect this assumption has on simplifying the wave equation.

The transformation

$$\psi = \rho^{-1/2}p \qquad (1.111)$$

puts (1.110) into the canonical form

$$(\nabla^2 + \omega^2/c^2)\psi = -\sigma\psi - \tau \ , \qquad (1.112)$$

where

$$\sigma = -3\left[\nabla\rho/(2\rho)\right]^2 + \nabla^2\rho/(2\rho) \qquad (1.113)$$

and

$$\tau = -\rho^{-3/2}\nabla\cdot(\rho\nabla\cdot\underline{\mathbf{F}}) \ . \qquad (1.114)$$

The significance of this form is that the induced source terms (¶1.3.3) on the right hand side of (1.112) are proportional to the field rather than to its derivatives.

1.3.5 Volume Source Formulation

For situations in which the acoustic parameters of the medium are unknown it is helpful to compare them with the known, constant parameters λ_o, μ_o, ρ_o of a reference medium. The wave velocity c_o in the reference medium is given by

$$c_o^2 = (\lambda_o + 2\mu_o)/\rho_o \ . \qquad (1.115)$$

Various indices may be defined to relate the two media. Wave velocities are compared using the refractive index n:

$$n = c_o/c \ . \qquad (1.116)$$

The wave number k is defined by

$$k = \omega/c_o \ . \qquad (1.117)$$

Relative variations in the constitutive parameters of the medium are defined by

$$\delta\lambda = (\lambda - \lambda_o)/(\lambda_o + 2\mu_o) \ , \qquad (1.118)$$

$$\delta\mu = (\mu - \mu_0)/(\lambda_0 + 2\mu_0) \ , \tag{1.119}$$

and

$$\delta\rho = (\rho - \rho_0)/\rho \ . \tag{1.120}$$

Using these definitions, the wave equation (1.101) may be written

$$\begin{aligned}(\nabla^2 + k^2)\nabla^2\phi = &- (\delta\lambda + 2\delta\mu)\nabla^2\phi - 2\nabla(\delta\lambda + 2\delta\mu)\cdot\nabla\nabla^2\phi \\ &- \nabla^2(\delta\lambda)\nabla^2\phi - k^2\nabla.(\delta\rho\nabla\phi) \\ &- \left[\nabla\nabla(2\delta\mu)\right]\cdot(\nabla\nabla\phi) - \kappa_0\nabla\cdot\nabla\cdot\underline{\mathbf{F}}(1 - \delta\lambda - \delta\mu)\end{aligned} \tag{1.121}$$

where terms of second order and higher in $\delta\mu$, $\delta\lambda$, and $\delta\rho$ have been dropped.

A comparison of (1.101) and (1.121) shows that the left hand side of (1.121) resembles that of (1.101) for the reference medium. The effects of inhomogeneities in the medium are described by the same terms on the left hand side of (1.121). The induced sources (¶1.3.3) are nonzero wherever the constitutive parameters change, which may be throughout $T_{\mathbf{r}}$. Thus (1.121) is referred to as the 'volume source' formulation of wave scattering (Bates and Ng 1972).

For an inviscid fluid the volume source formulation is obtained from (1.112). It is

$$(\nabla^2 + k^2)\psi = -q\psi - \tau \tag{1.122}$$

where

$$q = o + k^2(n^2 - 1) \ . \tag{1.123}$$

Many physical quantities obey (1.122). The new image reconstruction techniques developed in this thesis are based on (1.122) rather than (1.101) and so are applicable to many forms of imaging (Bates 1984).

1.3.6 Integral Representations of the Field

The wave equations, (1.112) and (1.122), may be transformed with the aid of Green's functions into integral representations for fields. Two alternatives exist depending on whether (1.112) or (1.122) is used.

A Green's function is, by definition, the field which emanates from a point source (Morse and Feshbach 1953 Ch. 7). Since an arbitrary source distribution can be thought of as a distribution of point sources, once the form of the Green's function for an arbitrarily positioned point source is known, the field from any distribution of sources can be formulated immediately. A point source at $\mathbf{r}'$, of unit strength, can be

represented by the three-dimensional delta function $\delta(\mathbf{r} - \mathbf{r}')$, which satisfies (Morse and Feshbach 1953 ¶2.1; Bracewell 1978 Ch. 5)

$$\delta(0) = \infty, \int_{T_{\mathbf{r}}} \delta(\mathbf{r} - \mathbf{r}')\, d\mathbf{r} = 1 \text{ and } \int_{T_{\bar{\mathbf{r}}}} \delta(\mathbf{r} - \mathbf{r}')\, d\mathbf{r} = 0 \tag{1.124}$$

where $\mathbf{r}$ is, and is not, a point of $T_{\mathbf{r}}$ and $T_{\bar{\mathbf{r}}}$ respectively. Since $d\mathbf{r}$ is a volume element, the physical dimensions of the three-dimensional delta function are $(\text{length})^{-3}$ implying that $\delta(\mathbf{r} - \mathbf{r}')$ has the physical dimensions of a density. It follows from (1.112) that the Green's function $G(\mathbf{r}|\mathbf{r}')$, describing the field at $\mathbf{r}$ due to a point source at $\mathbf{r}'$, must satisfy

$$(\nabla^2 + \omega^2/c^2)\, G(\mathbf{r}|\mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}') , \tag{1.125}$$

the solution of which is (Morse and Feshbach 1953 ¶7.2)

$$G(\mathbf{r}|\mathbf{r}') = [\exp(-jk|\mathbf{r} - \mathbf{r}'|)]/(4\pi|\mathbf{r} - \mathbf{r}'|) . \tag{1.126}$$

Now an arbitrary source distribution $S(\mathbf{r})$, occupying the region $\tilde{T}$, can be reexpressed with the aid of (1.124) as a distribution of point sources, i.e.

$$S(\mathbf{r}) = \int_{\tilde{T}} S(\mathbf{r}')\, \delta(\mathbf{r} - \mathbf{r}')\, d\mathbf{r}' . \tag{1.127}$$

When (1.127) is substituted into (1.112), and (1.124) through (1.126) are invoked, it is seen that

$$\psi(\mathbf{r}) = \int_{T_{\mathbf{r}}} [o(\mathbf{r})\, \psi(\mathbf{r}) + \tau(\mathbf{r})]\, G(\mathbf{r}|\mathbf{r}')\, d\mathbf{r}' . \tag{1.128}$$

When (1.127) is substituted into (1.122), it is similarly seen that

$$\psi(\mathbf{r}) = \int_{T_{\mathbf{r}}} [q(\mathbf{r})\, \psi(\mathbf{r}) + \tau(\mathbf{r})]\, G_0(\mathbf{r}|\mathbf{r}')\, d\mathbf{r}' \tag{1.129}$$

where $G_0(\mathbf{r}|\mathbf{r}')$ is the Green's function for the reference medium and satisfies

$$(\nabla^2 + k^2)\, G_0(\mathbf{r}|\mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}') . \tag{1.130}$$

In an imaging context, the representation (1.129) is of more immediate utility than (1.128). This is because the form of $G_0(\mathbf{r}|\mathbf{r}')$ is determined by the known reference medium while that of $G(\mathbf{r}|\mathbf{r}')$ is dependent on the velocity variations throughout the object which may not be known a priori.

1.4 DIRECT MEASUREMENT VERSUS REMOTE MEASUREMENT

1.4.1 Introduction

There are two approaches to reconstructing an image of the constitutive parameters of an object. Conceptually, the simplest is called 'direct' measurement (Finkelstein 1973). In this technique the field is measured throughout the region occupied by the object. The constitutive parameters for each location within the object are then calculated from these field measurements. The difficulties associated with making measurements within the object are discussed in ¶1.4.4. The second technique, called 'remote' measurement, utilises the propagation of fields to deduce the parameters from field measurements made some distance away from the object. In this section the two techniques are compared by considering a general sensing facility which incorporates the relevant features of both approaches.

1.4.2 The Sensing Facility

The problem which both direct and remote measurement attempt to solve may be called the 'measurement' problem. In this problem the constitutive parameters of an object at a measurement site M are to be made available at a reception point Q. Typically, a facility designed to solve this problem consists of three components, as illustrated in Fig. 1.2. Firstly, there is a probe device situated at a location P, called the 'probe site', which responds to the fields in the neighbourhood of P. The probe transmits the information it gathers through a communications link to a signal processor. The processor transforms the information obtained at P into the values of the object's parameters at M and then displays these values at Q. Data may need to be collected from several probe sites before the object's parameters can be calculated at the corresponding measurement sites. This is the case for Fourier imaging (Wolf 1969) and for computed-tomography (Bates et al. 1983).

Direct measurement is characterised by M and P being effectively identical. So, for measurements within the object, the probe must itself be within the object. For such measurements, part of the communications link must also be within the object. In remote measurement no part of the sensing facility is within the object. An example of the 'object' region which might be occupied by the object for each of the two techniques is shown in Fig. 1.2.

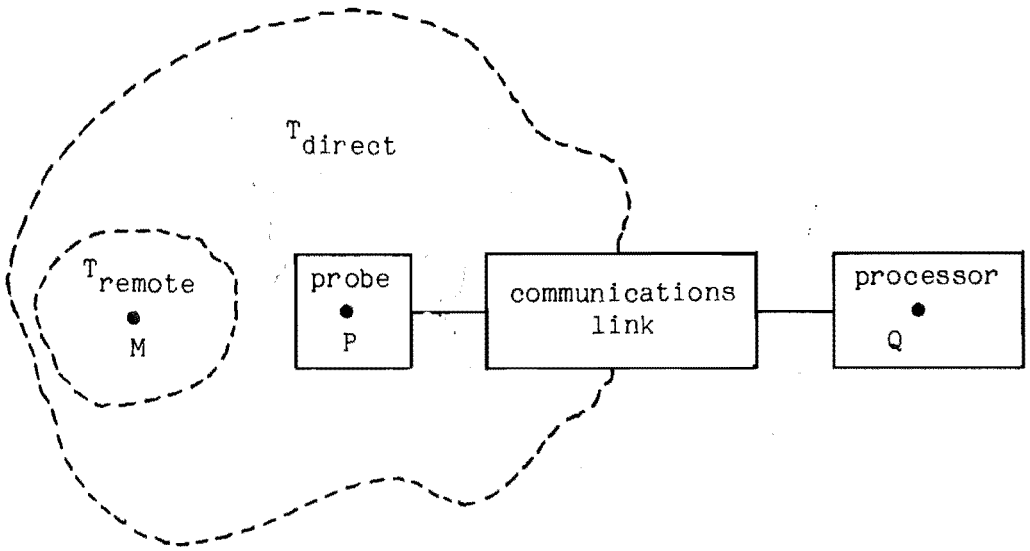


Fig. 1.2 The sensing facility. M = measurement site. P = probe site. Q = reception site. T_{remote} = object region for remote sensing. T_{direct} = object region for direct measurement.

1.4.3 Non-Interference

Ideally, the operation of the sensing facility should not disturb the object. This is impracticable, even in principle, since without some energy transfer between the object and the facility there can be no information transfer either.

The ideal may be approached in two ways. One way is to disturb the object but to return it to its initial state or to a state which is as close to that as possible. A somewhat unsubtle example of this 'reversible' measurement technique is that of explorative surgery. The object, in this case a patient, is cut open in order to observe some region and then sewn back up again. A more delicate example is provided by the ultrasonic imaging of a patient (Wells 1977 ¶5). In this case particles of tissue insonified by the beam of ultrasound are displaced from their equilibrium positions in a sinusoidal manner during the measurement but eventually return to their equilibrium positions.

An alternative to the reversible approach is to tolerate a permanent alteration but to ensure an that it is acceptably small. This can be called an 'irreversible' measurement. An example of one is the use of X-rays to measure tissue properties, since any non-zero intensity of X-radiation produces a non-zero probability of tissue damage (Mole 1978). This is in contrast to the effect of ultrasound, which is thought not to cause any permanent alterations to the tissue at the intensity levels used in diagnostic medicine (Sokollu 1971). At higher intensity levels of ultrasound, however, tissue damage does occur (Wells 1977 ¶10.2).

1.4.4 Economics

As with any facility, the costs associated with sensing may be divided up into two categories. In the first category one has the fixed purchase costs for each component in the facility, while in the second there are the running costs incurred each time the facility is used. The fixed costs are influenced by the enviroment in which the equipment must operate. In this area, direct measurement is usually at a disadvantage with respect to remote measurement since two of the components of the sensing facility must be capable of operating within the object. These are the probe and the communications link. As an example of how enviromental conditions can effect the design and therefore the cost of these components, consider the use of a catheter to monitor activity within the heart (Cromwell et al. 1980). The catheter must be small enough not to interfere with blood flow in the vessels in which it is placed, and it must be made of inert materials. Both of these requirements combine to make a catheter an expensive item of equipment. In some situations such as the imaging of an X-ray emissivity distribution in a plasma confinement device, as described by Chase et al. (1981), the object constitutes such a hostile enviroment that equipment cannot easily be designed to operate within it and remote measurement must be employed. Direct measurement is also at a disadvantage with respect to running costs. This is because the probe and the communications link must be introduced into each object that is to be imaged and it is usually more difficult to move the probe around within the object than it would be if it were in free space. In remote measurement the probe can be moved around in free space without interfering with the object. For catheterisation, the running costs are high because a number of skilled people must be in attendance as a vein is opened up to allow the catheter to be inserted into the patient and for the subsequent procedures.

Although the probe and the communications link may be less expensive for remote measurement, additional processing is required to deduce the object parameters at M from the measurements at P. For imaging based on echo-location principles this processing is quite simple and need not be expensive to implement. This is not necessarily the case for more complicated methods of image reconstruction, however. For methods such as Fourier imaging and computed-tomography a fast and versatile computer is required.

1.4.5 Convenience

When the object being imaged is a human there are two additional factors which need to be considered. These factors are patient safety and comfort. The non-invasive methods of remote measurement are more comfortable since a surgical operation and the consequent recovery period are avoided. Remote measurement methods are also quicker and this contributes to patient comfort. The avoidance of surgery also contributes to patient safety.

1.5 THE REMOTE PROBING PROBLEM

1.5.1 Components of the Physical System

In this section the remote probing problem is defined with the aid of the wave equation

$$(\nabla^2 + k^2)\psi = -q\psi + s \tag{1.131}$$

where q is referred to as the 'scattering potential' and s is the (free) source term. Longitudinal waves in inviscid fluids are described by (1.131) (see (1.122)). The triple (ψ, q, s) is called the 'physical system'. Four subsystems of the physical system need to be identified to allow the remote probing problem to be defined. A similar approach is employed by Bates and McKinnon (1980).

The object to be probed is a subsystem of the physical system and is described by the triple $(T_{obj}, q_{obj}, s_{obj})$. The components of the triple are the volume occupied by the object, the induced sources, and the free sources respectively. A subsystem $(T_{gen}, q_{gen}, s_{gen})$ is used to generate a wave which impinges upon the object. The wave scattered by the object is measured by a subsystem $(T_{meas}, q_{meas}, s_{meas})$. All portions of the physical system not included in these subsystems are included in a homogeneous subsystem (T_o, q_o, s_o). Fig. 1.3 illustrates these subsystems.

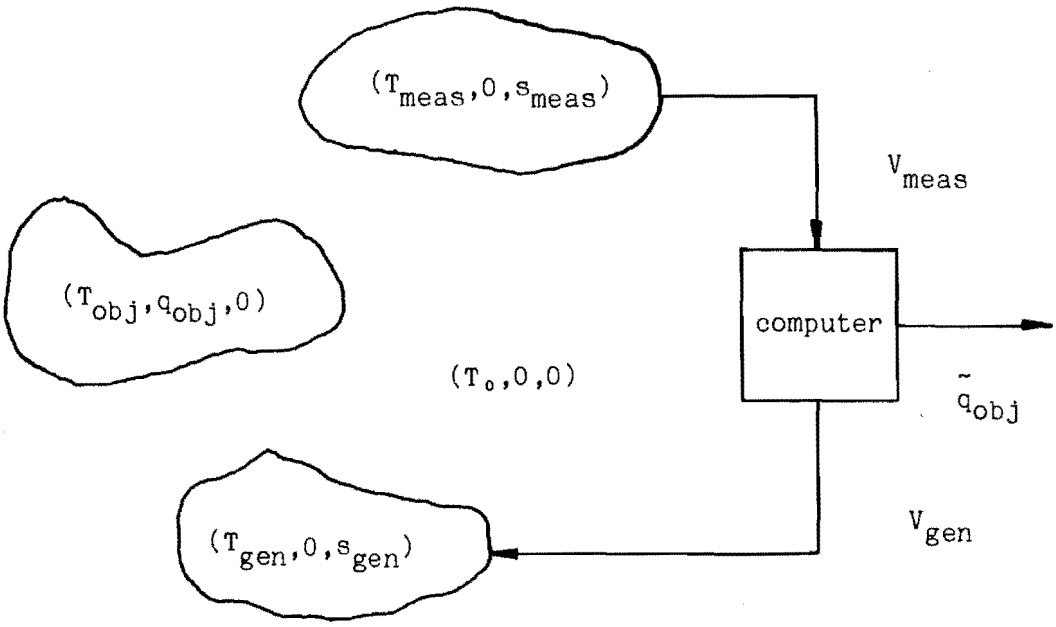


Fig. 1.3 The physical system and the computer for CAI.

Free sources are assumed to exist only in the generation and measurement regions,

$$s_{obj} = s_o = 0 \ . \tag{1.132}$$

In the homogeneous region,

$$q_o = 0 \ , \tag{1.133}$$

by definition. Also, for ideal generation and measurement,

$$q_{meas} = q_{gen} = 0 \ . \tag{1.134}$$

The total system is the union of these subsystems,

$$T_{\mathbf{r}} = T_{gen} \ U \ T_o \ U \ T_{obj} \ U \ T_{meas} \ , \tag{1.135}$$

$$q = q_{obj} \ , \tag{1.136}$$

$$s = s_{gen} + s_{meas} \ . \tag{1.137}$$

The total field may also be expressed as a sum with the aid of (1.129),

$$\psi = \psi_{gen} + \psi_{obj} + \psi_{meas} , \tag{1.138}$$

where

$$\psi_{gen}(\mathbf{r}) = \int_{T_{gen}} s_{gen}(\mathbf{r}') \, G_o(\mathbf{r}|\mathbf{r}') \, d\mathbf{r}' , \tag{1.139}$$

$$\psi_{obj}(\mathbf{r}) = \int_{T_{obj}} q_{obj}(\mathbf{r}') \, \psi(\mathbf{r}') \, G_o(\mathbf{r}|\mathbf{r}') \, d\mathbf{r}' , \tag{1.140}$$

$$\psi_{meas}(\mathbf{r}) = \int_{T_{meas}} s_{meas}(\mathbf{r}') \, G_o(\mathbf{r}|\mathbf{r}') \, d\mathbf{r}' . \tag{1.141}$$

When the sources s_{gen} are under external control, the field ψ_{gen} is known since there are no induced sources in T_{gen}.

1.5.2 The Data Available from the System

The data available from the physical system consists of a finite number M of source pairs. The first member of the m^{th} pair is the applied source $s_{gen,m}$ and the second is the source $s_{meas,m}$ induced in T_{meas} when the source in the generation region is $s_{gen,m}$. The set D of all the pairs is written as

$$D = \{(s_{gen,1}, s_{meas,1}), \ldots, (s_{gen,M}, s_{meas,M})\} . \tag{1.142}$$

1.5.3 Definition of the Problem

The remote probing problem is to find an estimate or image $\hat{q}_{obj}$ of q_{obj} from the measurement pairs D, under the restriction

$$(T_{gen} \cup T_{meas}) \cap T_{obj} = \emptyset \tag{1.143}$$

where $\emptyset$ is the empty set. (1.143) simply says that the generation and measurement regions should be remote from the object. The region occupied by the object may or may not be known. When the sources s_{gen} are known but not controlled, the problem is that of remote sensing. The differences between the estimate $\hat{q}_{obj}$ and the object q_{obj} are called 'artefacts'.

1.5.4 Point Spread Invariant Systems

For some CAI systems the estimate $\hat{q}_{obj}$ is related to the object q_{obj} by the relationship:

$$\hat{q}_{obj} = h \odot q_{obj} + c \tag{1.144}$$

where h and c are functions which are independent of the nature of the object. Such systems have a very important property. Consider an object f consisting of a δ function located at $\mathbf{r}$. This object may be written as

$$f = T_{\mathbf{r}}(\delta) \tag{1.145}$$

with the aid of the translation operator defined by (1.31). The image $\hat{f}$ of f is found by substituting (1.145) into (1.144) and invoking the identity (1.45), i.e.

$$\begin{aligned} \hat{f} &= h \odot f + c \\ &= h \odot T_{\mathbf{r}}(\delta) + c \\ &= T_{\mathbf{r}}(h \odot \delta) + c \\ &= T_{\mathbf{r}}(h) + c \ . \end{aligned} \tag{1.146}$$

Thus the image consists of two parts. The first is the function h shifted to the same location as the δ function. The second part is the function c. The function h is therefore called the 'point spread' function (PSF) of the system. Any system which is described by (1.144) is said to be point spread invariant (PSI) (Bates and McDonnell 1985 ¶3). The function c is called the 'contamination'. Both functions can have a stochastic and a deterministic component. Approximations made in the derivation of the reconstruction algorithm employed by the CAI system are usually responsible for the deterministic component. Measurement noise is usually represented by a stochastic function. While in theory h and c are independent of the object, it is sometimes helpful to model CAI systems for which this is not true by (1.144). This is discussed further in ¶6.8. For a PSI CAI system to be useful, the contamination must be small with respect to the 'isoplanatic component' $h \odot q_{obj}$. The region in which the PSF is significantly different from zero is called the 'resolution cell'.

1.5.5 Isolating Constitutive Parameters

The estimate obtained from remote probing may only partially characterise the object. This is so if all of the source pairs are obtained at the same frequency, as can be seen from an examination of (1.123). Refractive index and density variations both give rise to scattering but they are lumped together in the single induced source distribution q.

The differing frequency dependences of the two contributions may allow the two to be separated, provided that measurements are made at more than one frequency (Nahamoo and Kak 1982 Ch. 4).

To illustrate the difficulties encountered when the constitutive parameters are to be isolated from q_{obj}, consider the special case of ultrasonic imaging of an inviscid fluid. The two parameters to be imaged are the compressibility and the density. For some fluid models the two parameters cannot be separated, even in principle. If the object is modelled as a region of variable density, for instance, (1.109) shows that at best only the product of the two parameters can be inferred.

Even when both parameters appear independently in the model, the data set may not enable the parameters to be imaged independently. As an example, consider the model (1.112). Gore and Leeman (1977) investigate the use of this model for the analysis of 'pulse-echo' measurements as obtained from conventional medical ultrasonic imaging systems. In the terminology of ¶1.5.2 a pulse-echo measurement is one for which

$$s_{gen,m} = \delta(\mathbf{r} - \mathbf{r}_m) \, s_{gen}(t) \tag{1.147}$$

and

$$s_{meas,m} = \delta(\mathbf{r} - \mathbf{r}_m) \, s_{meas}(t) \tag{1.148}$$

where $s_{gen}(t)$ is a generated pulse, $s_{meas}(t)$ is the echo received and generation and reception take place at the same point $\mathbf{r}_m$. They find that the measurements are of the difference between the compressibility variation and the density variation, integrated over a surface determined by $\mathbf{r}_m$. For the idealised point sources of (1.147) the integration is over the surface of a sphere. The density variation $\delta\rho$ is defined by (1.120) and the compressibility variation $\delta\kappa$ is defined by

$$\delta\kappa = (\kappa - \kappa_0)/\kappa \, . \tag{1.149}$$

To first order,

$$\delta\kappa = -\delta\lambda \tag{1.150}$$

where $\delta\lambda$ is defined, in a manner analogous to (1.149), by (1.118). Norton and Linzer (1981) define a scalar reflectivity function,

$$f(\mathbf{r}) = \delta\kappa(\mathbf{r}) - \delta\rho(\mathbf{r}), \tag{1.151}$$

and show how it may be reconstructed from pulse-echo data. Norton (1983) proposes a method for separating the terms in (1.151). The basis of the method is the observation that a localised compressibility variation

scatters like a monopole and that a localised density variation scatters like a dipole (Morse and Ingard 1968 p.410). In theory, therefore, the two components can be distinguished by their different angular dependences. Compressibility scattering is isotropic while density scattering has a $\cos(\theta)$ dependence, where θ is the angle between two direction vectors, one pointing at the generation site and the other toward the reception site. The reflectivity defined by (1.151) is modified to

$$f(\mathbf{r},\theta) = \delta\kappa(\mathbf{r}) - \delta\rho(\mathbf{r}) \cos(\theta) \ . \qquad (1.152)$$

Norton's method uses data from back scattering ($\theta = 0$) and for $\theta = \pi/2$ to isolate the two terms of (1.152). In practice, high resolution imaging of compressibility and density has yet to be achieved. Part of the reason for this is the failure of the Born approximation, which is assumed by all of the workers referenced in this section. The Born approximation is discussed in ¶1.6.3.

1.5.6 Computer-Aided Imaging

A computer of some sort is required to calculate the image $\hat{q}_{obj}$ from the data D. This computer may be very simple, e.g. a lens, or very complicated, e.g. an electronic general purpose digital computer. Despite the diversity amongst imaging systems it is possible to identify features common to all of them. These features stem from the essential interface requirements of any imaging system. An input is required for s_{meas} and an output for $\hat{q}obj$. For remote probing there will also be an output for s_{gen}. Also, the emanation used by the computer for internal information transfer must be coupled to the appropriate external emanations at these interfaces.

1.5.7 Transduction

The Need for Transduction and Definition of Terms

In the CAI system described in ¶1.5.1 energy propagates in many forms. In the physical system, the field ψ is responsible for the transport of energy. A different form of energy may be used by the computer to carry out computations. Also, the port through which the recipient receives information may operate using yet another form of energy. If information is to be exchanged, the various forms of energy must be coupled at the interfaces between those systems. A device which does this is known as a 'transducer'. The process of converting one form of energy into another is called 'transduction'.

An Example

Consider an ultrasonic imaging system incorporating an electronic digital computer, and suppose that the reconstructed images are displayed upon a video monitor for inspection by a human operator. Transducers are used at three distinct places in this system. Electro-acoustic transducers are used to convert a voltage generated by the computer into an acoustic source s_{gen} (¶1.5.1) and to generate a voltage related to the induced source s_{meas}. The video monitor is an electro-optical transducer since it transforms electrical signals from the computer into an optical field. This is the appropriate form for the port, which is the human eye in this example.

The Signal-Dimension of a Transducer

A video monitor is an example of a complicated transducer. Electrical energy is converted to light on the surface of the screen. The rates of energy conversion at two points on the screen can be controlled independently, even when the points are quite close. In general, the 'signal-dimension' of a transducer can be defined to be the dimension of the signal which it is converting. For the video monitor, the signal-dimension is three, since there are two space dimensions and one time dimension.

The signal-dimension is distinct from the dimension of the region in which energy conversion takes place. For example, consider an electro-acoustic transducer to which only two electrodes are attached. The electrical input to this transducer is completely specified by a function of time only. At any instant, only one number is available to characterise the acoustic sources in the region occupied by the transducer. The electrical signal must therefore correspond to some form of average over space of the sources. Since the acoustic field at a point is given by one number, such a transducer is of signal dimension one and is called a 'point' transducer. However, the surface over which energy conversion occurs may be two-dimensional, e.g. a disc.

Impedance

A point transducer may be represented by a system which relates two pairs of variables. One member of each pair is like a force, the other is like a velocity. For an electro-acoustic transducer, one of the pairs consists of the voltage (force-like) and current (velocity-like). When

both of these are sinusoids of the same frequency, they may be related by a complex 'impedance'. The dependence of the impedance on frequency characterises the electrical response of the system.

In a similar manner the acoustic field is described by the pressure and the particle velocity. In this case the pressure is the force-like quantity. The acoustic impedance of the homogeneous reference medium (¶1.3.5) is (Wells 1977 ¶1.6)

$$Z_0 = \rho_0 c_0 \ . \qquad (1.153)$$

Efficiency and Coupling

The ability of a transducer to convert from one energy form to another, called the 'output' energy form, is measured by its coupling coefficient. The square of this coefficient is defined to be the ratio of the energy stored in the output energy form to the total energy stored within the transducer (Wells 1977 ¶3.1a). This value is distinct from the efficiency of the transducer, which measures the total amount of energy which goes into storage relative to the applied energy. The remainder of the applied energy is dissipated as heat.

An Overview of Transducer-Object Systems Analysis

Transform methods may be used to analyse the response of an electro-acoustic transducer in the presence of a scatterer. Kerns (1975) investigates a system in which the transducer is coupled to an elastic scatterer via a fluid bath. Pulse-echo measurements (¶1.5.5) are made, so the transducer is operating firstly in transmission mode and then in reception mode. His work is specialised by Johnson and Devaney (1982) for the case of a circular symmetric transducer and a stratified elastic half-space.

To perform the analysis, Johnson and Devaney (1982) express the acoustic field as a sum of plane waves (Clemmow 1966), and describe the electrical waveforms by their temporal Fourier transforms. Each plane wave is indexed by a parameter called the 'wave vector'. Each component of the system is characterised by a response function which relates output to input for a particular temporal frequency and/or wave vector. The components of the system which are considered are the generation and reception at the transducer, acoustic wave propagation to and from the scatterer, and the scattering from the object. They show that the electrical input-output response is in the form of a one-dimensional

integral transform of the reflection coefficient for the half-space. The transform is over angles of incidence. They also show that the integral transform can be approximated by the product of a transducer-dependent response function and the reflection coefficient for normal incidence. This result is equivalent to the 'lumped parameter' model proposed by Seki et al. (1956) and developed by Papadakis (1975).

The Radiation Pattern

The angular distribution of energy radiated from a transducer is referred to as the 'radiation pattern' of the transducer. Typically, it has a lobe structure. The solid angle into which most of the energy is transmitted is known as the 'main' lobe. Radiations of significant energy into other directions are called 'side' lobes.

Piezoelectricity

For certain materials an electric field is associated with an external stress,

$$F_{ij} = \sum_{k=1}^{3} -e_{kij} E_k \tag{1.154}$$

where **E** is the electric field vector and $\underline{\mathbf{e}}$ is the third rank piezoelectric tensor for the material (Sapriel 1979). The corresponding free source term is found by substituting (1.154) into (1.114) to be

$$\tau = -\frac{1}{\rho^{3/2}} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \frac{\partial}{\partial x_i} \left[\rho \frac{\partial}{\partial x_j} (e_{kij} E_k) \right] . \tag{1.155}$$

The link between transduction and the data set (1.142) is established by (1.155) when τ is identified with s of (1.131).

1.6 SOLUTION TECHNIQUES FOR SCATTERING PROBLEMS

1.6.1 Direct Methods for Solving Integral Equations

The calculation of the field scattered by an object from knowledge of the object and the incident field is known as the 'forward scattering' problem. In this section methods for solving the forward scattering problem are discussed.

When the total field within the object is known, the field scattered by the object may be calculated using (1.140). However, since the total field within the object is due in part to ψ_{obj}, the solution of (1.140), given q_{obj} and ψ_{gen}, cannot usually be obtained simply. An exact solution is possible for some homogeneous objects of simple shape (Bowman et al. 1969). For inhomogeneous objects or objects with complicated boundaries, approximate methods of solution are usually required.

Bates (1975) reviews methods for electromagnetic field computations. In particular the practical computational significance of various analytic constraints is considered. The methods for solving the scattering problem are partitioned into three classes. In the 'differential' approach the wave equation is solved directly by such techniques as 'finite difference' (Kunz and Lee 1978a, 1978b; Mur 1981a, 1981b) or 'finite element' (Silvester 1969; Daly 1974) methods. Methods in which the field is expanded in sequences of functions which are proper solutions of the wave equations comprise the 'series' approach. An example of the series approach is 'point-matching', wherein the series coefficients are determined by satisfying the boundary conditions at a number of points. Thus method is discussed by Bates (1967), Lewin (1970a, 1970b) and Bates et al. (1973). The third class, referred to as the 'integral' approach, is based upon integral equation formulations such as (1.140).

(1.140) is a Fredholm integral equation of the second kind for ψ_{obj} or ψ (Morse and Feshbach 1953 Ch. 8; Pogorzelski 1966 Ch. 2) and may be written in operator form as

$$\psi_{obj} = \psi - \psi_{gen} = L(\psi) \tag{1.156}$$

where

$$L(\psi) = \int_{T_{obj}} q_{obj}(\mathbf{r}')\, \psi(\mathbf{r}')\, G_0(\mathbf{r}|\mathbf{r}')\, d\mathbf{r}' \tag{1.157}$$

is a linear integral operator, and the field ψ_{meas} is neglected.

A straightforward technique for solving (1.156) is to express ψ as a linear combination of basis functions (Green 1969 ¶4.2),

$$\psi = \sum_m a_m \psi_m \tag{1.158}$$

where $\{\psi_m\}$ is a set of basis functions, referred to as the 'basis', and the a_m are constants. Since L is a linear operator its effect on an arbitrary wave can be computed in terms of its effect on each member of the basis. The result of applying L to ψ_m may also be expanded in terms of

the basis,

$$L(\psi_m) = \sum_n L_{nm}\ \psi_m \tag{1.159}$$

where the L_{nm} are constants. When the expansions (1.159) and (1.158) are substituted into (1.156), the integral equation is replaced by a system of simultaneous linear equations,

$$b_m = \sum_n (\delta_{mn} - L_{mn})\ a_n \tag{1.160}$$

where the b_m are the expansion coefficients of the incident field

$$\psi_{gen} = \sum_m b_m\ \psi_m \ . \tag{1.161}$$

In order for the object to be represented adequately by the constants L_{nm}, one would expect the number of these constants to be of the same order of magnitude as the number of resolution cells in the object. Since the scattered wave can be expected to vary appreciably over a wavelength, the linear dimensions of a resolution cell are chosen to be one order of magnitude less than a wavelength. Now since the number of operations required to invert the $N \times N$ matrix $\delta_{nm} - L_{nm}$ is of order N^3 (Nash 1979 p.40), the numerical solution of (1.160) rapidly becomes impractical as the size of the object relative to the wavelength increases.

The 'method of moments' (Harrington 1968) is a technique for solving linear operator equations such as (1.156) which is more general than the method described by (1.158) through (1.161). The kernel of (1.140) is singular and this can cause difficulties in its numerical solution. Also there is a uniqueness problem because of cavity resonances for totally reflecting objects. These two difficulties are overcome in the 'null-field' approach which replaces (1.131) by an infinite set of non-singular integral equations (Bates and Wall 1977a, 1977b, 1977c). A surface integral representation, suitable for numerical computation of electromagnetic scattering by dielectric cylinders, is presented by Morita (1978).

1.6.2 Approximate Methods

To overcome the practical difficulties of solving (1.160), approximate solutions of (1.156) are used (Bates 1969). Two of the principal methods employed are high frequency asymptotic solutions and perturbation techniques. The latter are appropriate when the object is tenuous and consequently the scattering is weak. These methods are discussed further in ¶1.6.3.

The high frequency methods use ray techniques. They are applied by subdividing the object up into components for which canonical solutions exist. In order of increasing accuracy the methods are: 'geometrical optics' (Born and Wolf 1965), the 'geometrical theory of diffraction' (Keller et al. 1956; Kouyoumjian 1965; James 1976; Seckler and Keller 1959) and the 'uniform asymptotic theory' (Lee 1978). The discovery of the geometrical theory of diffraction has prompted the reexamination of the theory of geometrical optics. Cornbleet (1983) reviews the development of geometrical optics from a design tool for optical systems to an approximate theory of the electromagnetic field.

The methods of 'physical optics' (Jones 1964 pp.633-640; Bates 1974) and the 'physical theory of diffraction' (Ufimstev 1962) are based on replacing the object with the sources generated by the incident wave. The scattering is then calculated from the radiation of these induced sources.

Ursell (1966) shows that physical optics is exact for smooth convex bodies in the limit of infinitely high frequencies and may thus be classed as a high frequency technique. The physical optics solution for an infinite plane is exact at all frequencies. Physical optics is applicable to objects which have well defined boundaries. Although the method is particularly well suited to totally reflecting objects, source densities can usefully be postulated on the surfaces of penetrable bodies (Bechmann and Spizzichimo 1963 Ch. 3). Bates and Wall (1977b) describe a generalised physical optics which is exact for certain scatterers in addition to the infinite plane. Physical optics is able to predict some diffraction effects and is therefore more accurate than geometrical optics. The geometrical theory of diffraction is more accurate still and gives predictions which are only slightly less accurate than the physical theory of diffraction.

1.6.3 Perturbation Methods

Perturbation methods for calculating the scattering are derived from an analytic technique for solving integral equations called the 'method of successive approximations' (Zabreyko et al. 1975 ¶2.1). In this method the solution is expressed as a series. The convergence of the series is determined by the value of the largest eigenvalue of L. If this eigenvalue is less than one the series converges. In practice, it is difficult to verify this condition although at sufficiently low frequencies

it can be shown that this condition holds (Kleinman 1965). Sufficiently low, in this case, means that the wavelength is long compared to the scatterer. This scattering region is called the 'Rayleigh' region. It is of limited relevance to the remote probing problem since the object to be imaged is usually many wavelengths in across.

The method of successive approximations is used to solve (1.156) by writing the total field as a series,

$$\psi = \psi_{gen} + \sum_{m=1}^{\infty} \psi_m^{(B)} \tag{1.162}$$

and calculating $\psi_m^{(B)}$ from $\psi_{m-1}^{(B)}$ using the formula

$$\psi_m^{(B)} = L(\psi_{m-1}^{(B)}) \tag{1.163}$$

where $\psi_o^{(B)}$ is defined to be the incident field,

$$\psi_o^{(B)} = \psi_{gen} \ . \tag{1.164}$$

Physically, (1.162) to (1.164) express the solution in terms of multiple scattering. First order scattering $\psi_1^{(B)}$ results from the interaction of the incident wave and the object. When $\psi_1^{(B)}$ interacts with the object, second order scattering $\psi_2^{(B)}$ is produced and so on. When the series (1.162) converges it usually does so very slowly unless q_{obj} is small (Morse and Ingard 1968 ¶8.1; Cowley 1975 ¶1.5). In this case, the first and second order terms of (1.162) account for most of the total field in any case. The sum of the first two terms of (1.162) is known as the 'first Born' approximation $\psi^{(B)}$ to the total field,

$$\psi^{(B)} = \psi_{gen} + \psi_1^{(B)} \ . \tag{1.165}$$

An alternative to the Born approximation, called the 'Rytov' approximation, results when the successive approximation method is applied to a transformed version of the wave equation. The transformation employed is

$$\psi = \exp(j\phi) \tag{1.166}$$

where the field ψ is expressed in terms of a complex phase ϕ. When (1.166) is substituted into the wave equation (1.122), a Riccati equation

$$\nabla\cdot(\nabla\phi) + \nabla\phi\cdot\nabla\phi + k^2 = -q_{obj} \tag{1.167}$$

for $\nabla\phi$ is obtained (Ishimaru 1978 ¶17.2.2). As for the Born approximation, the method of successive approximations is applied by writing the phase ϕ as a series,

$$\phi = \phi_{gen} + \sum_{m=1}^{\infty} \phi_m^{(R)} \tag{1.168}$$

where ϕ_{gen} is defined by

$$\psi_{gen} = \exp[j\phi_{gen}] \ . \tag{1.169}$$

The formulas for calculating the $\phi_m^{(R)}$ are

$$\phi_1^{(R)}(\mathbf{r}) = \frac{1}{k^2\psi_{gen}} \int_{T_{obj}} q_{obj}(\mathbf{r}')\ \psi_{gen}(\mathbf{r}')\ G_0(\mathbf{r}|\mathbf{r}')\ d\mathbf{r}' \tag{1.170}$$

and

$$\phi_m^{(R)}(\mathbf{r}) = \frac{1}{\psi_{gen}} \int_{T_{obj}} |\nabla'\phi_{m-1}^{(R)}(\mathbf{r}')|^2\ \psi_{gen}(\mathbf{r}')\ G_0(\mathbf{r}|\mathbf{r}')\ d\mathbf{r}' \tag{1.171}$$

where ∇' denotes a gradient operator which acts with respect to the point $\mathbf{r}'$. The first Rytov approximation $\psi^{(R)}$ for the field ψ is defined by

$$\psi^{(R)} = \exp[\phi_{gen} + \phi_1^{(R)}] \ . \tag{1.172}$$

This is related to the Born approximation by (Ishimaru 1978 ¶6.6)

$$\psi^{(R)} = \psi_{gen}^{(B)} \exp[\psi_1^{(B)}/\psi_{gen}] \ . \tag{1.173}$$

Chernov (1960) specifies the conditions under which the approximations are valid. For the Born approximation, the first order scattering ψ_1 must satisfy

$$|\psi_1^{(B)}/\psi_{gen}| << 1 \tag{1.174}$$

so that the relative scattering at all points within the object must be small. The condition for the validity of the Rytov approximation is

$$|\nabla(\phi_1^{(R)})|/k << 1 \tag{1.175}$$

and so the relative scattering per wavelength must be small.

Keller (1969) considers the scattering of a plane wave in a situation which has an exact solution. The Born and Rytov approximations are compared by determining the dependence of the error on the parameters of the scattering problem. These parameters are the extent of the scattering region and the amplitude of the inhomogeneity. It is shown that both methods are equally accurate in terms of the latter but that the Rytov method is valid for larger scattering regions than the Born approximation,

as expected from (1.174) and (1.175). However, the superiority of the Rytov method is dependent on the assumption of a single plane wave. In situations where many plane waves are present, the Rytov approximation must be applied to each separately, if its advantages are to realised.

When the scattering is predominantly in the forward direction, the wave propagation is along a ray path which is in general curved. Bates et al. (1976) describe a method called the 'extended' Rytov approximation which partially takes into account the $|\nabla'\phi_{m-1}^{(R)}(\mathbf{r}')|^2$ factor in the integrand of (1.171) for m = 2. The method assumes that the total field within the object propagates along rays which are only slightly curved. Numerical results presented by Dunlop et al. (1976) suggest that the method is more accurate than the Rytov and Born approximations for scattering in the forward direction.

Ishimaru (1978) argues that the Rytov approximation is appropriate for problems where only scattering in the forward direction is of interest while the Born approximation is more suitable when the scattering in all directions needs to be calculated. To illustrate the significance of the Rytov approximation (1.169) in the forward direction, consider a simple model of the object. It is taken to consist of a sequence of slabs for which all the interfaces are parallel to the yz plane. The generated field is assumed to be propagating in the x direction. The i^{th} slab, counting along the x axis, extends between x_{i-1} and x_i, where the first interface is at x_o. Assuming that the transmission coefficient at the i^{th} interface is T_i, the field $\psi_i^{(T)}$ between x_{i-1} and x_i is given by

$$\psi_i^{(T)} = \psi_{gen} \prod_{m=0}^{i-1} T_m \ . \tag{1.176}$$

The Rytov transformation may be applied to the transmission coefficients

$$T_m = \exp[\phi_m^{(T)}] \tag{1.177}$$

and to the fields

$$\psi_i^{(T)} = \exp[\phi_i^{(T)}] \ . \tag{1.178}$$

When (1.169), (1.177), and (1.178) are substituted into (1.176) the result is

$$\phi_i^{(T)} = \phi_{gen} + \sum_{m=0}^{i-1} \phi_m^{(T)} \ . \tag{1.179}$$

Hence, the Rytov transformation converts the product of (1.176) into a sum

where each term corresponds to an inhomogeneity of the object. The integral equation (1.140), upon which the Born approximation is based, can also be considered as a sum of inhomogeneities. However, in the formulation of (1.179), the total phase shift may be appreciable even though each of the inhomogeneities is small. Thus, the Rytov approximation is the natural one to use for scattering in the forward direction.

A comparison of the Born and Rytov approximations in the context of acoustic CAI is given by Kaveh et al. (1982).

1.7 SOLUTION OF THE REMOTE PROBING PROBLEM BY FOURIER IMAGING

In this section solutions to the remote probing problem are derived. The solutions are based on the perturbation methods discussed in ¶1.6.3 and are derived within the first Born and Rytov approximations. In both cases second and higher order scattering are neglected. Hence, the solutions are applicable only to the remote probing of weakly scattering objects.

Two classes of data D (¶1.5.2) are considered. The two classes differ only with respect to the region in which measurements are made. The relative locations of the two measurement regions are illustrated in Fig. 1.4. For both classes the applied sources $s_{gen,m}$ in T_{gen} are such that the generated field ψ_{gen} is a plane wave

$$\psi_{gen}(\mathbf{r},k) = A_{gen}(k)\ \exp(-j\mathbf{k}\cdot\mathbf{r}) \tag{1.180}$$

with wave vector $\mathbf{k}$ and amplitude $A_{gen}(k)$. In practice, a spectrum of plane waves with the same unit wave vector $\hat{\mathbf{k}}$ but differing wave numbers k, may be applied simultaneously. A narrow pulse and a frequency chirp are examples of such insonifications.

The first class is considered in ¶1.7.1 and consists of 'projection' measurements. These are measurements which are made along a line perpendicular to the wave vector. The region containing the line is denoted $T_{meas}^{(P)}$.

The set of locations for $T_{meas}^{(P)}$ subject to (1.143) may be naturally partitioned into two subsets. A typical member of one of the subsets is located on the side of the object opposite to the applied sources and is denoted by $T_{meas}^{(PF)}$. A member of the other set is located on the same side and is denoted by $T_{meas}^{(PB)}$.

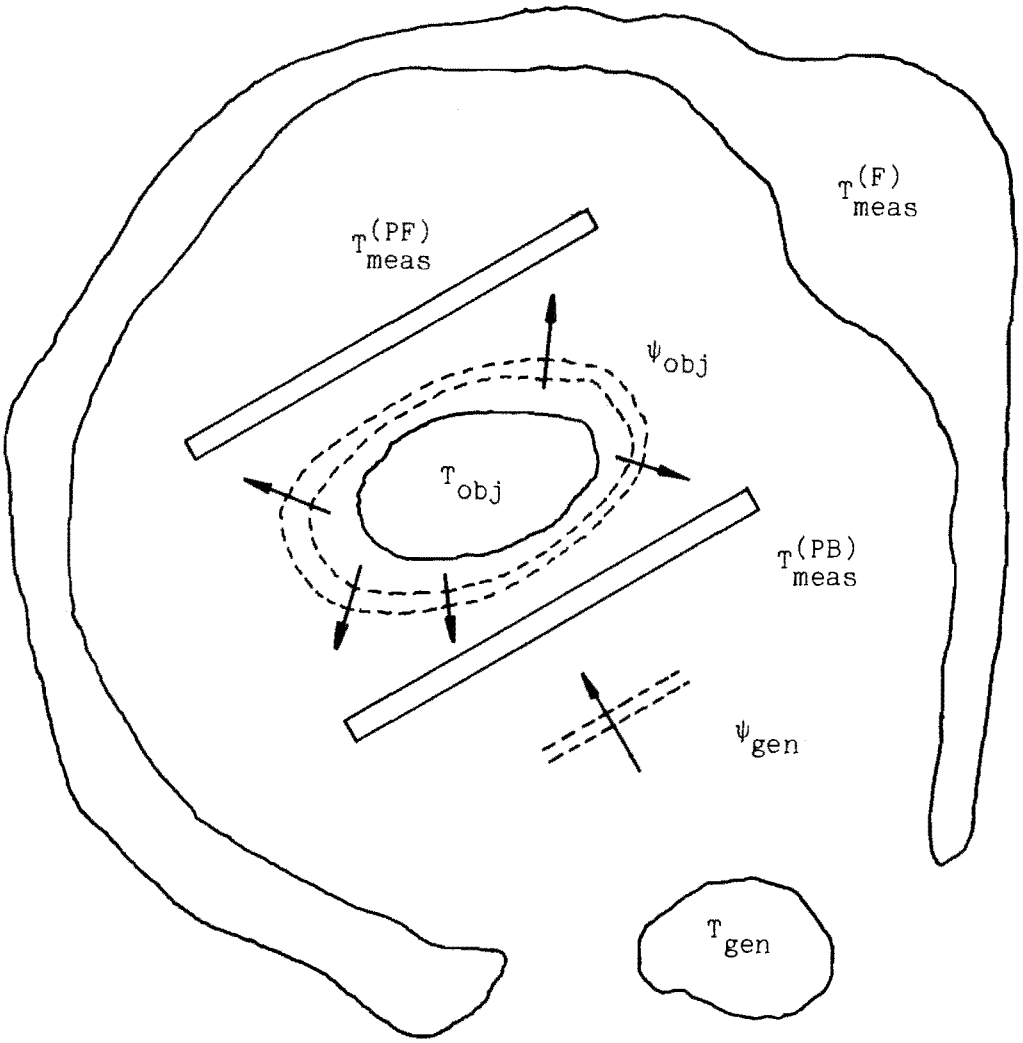


Fig. 1.4 Two possible locations for T_{meas}. The object residing in the region T_{obj} is insonified by a plane wave ψ_{gen} generated by sources within T_{gen}. For projection measurements, the wave scattered by the object ψ_{obj} is measured on a straight line within $T^{(PF)}_{meas}$ or $T^{(PB)}_{meas}$. For far-field measurements, ψ_{obj} is measured within $T^{(F)}_{meas}$.

An especially important feature of $T_{meas}^{(P)}$ is that it may be arbitrarily close to T_{obj}.

Measurements made in a region $T_{meas}^{(F)}$ some distance from T_{obj} belong to the second class which is referred to as the 'far-field' measurement class. This class is discussed in ¶1.7.2 and the allowable locations for $T_{meas}^{(F)}$ are derived there. In contrast to $T_{meas}^{(P)}$, $T_{meas}^{(F)}$ may sense scattering from the object in all directions. The implications of this for object reconstruction are discussed in ¶1.7.3.

It is shown in ¶1.7.1 and ¶1.7.2 that a relation of the form

$$\chi(\mathbf{u}) = \tilde{q}_{obj}(\mathbf{u}) \quad \text{for } \mathbf{u} \text{ in } T_c \tag{1.181}$$

exists for both classes, where χ is a function of the measurements and $\tilde{q}_{obj}$ is the two-dimensional Fourier transform (1.56) of the object image q_{obj}. For a single wave vector, T_c is an arc in Fourier space $T_{\mathbf{u}}$ and so χ is a slice of $\tilde{q}_{obj}$ on T_c. The equation (1.181) is therefore known as the 'generalised projection theorem'. In view of the occurrence of the Fourier transform of the object on the right hand side of (1.181), all solutions to the remote probing problem based on (1.181) are referred to as examples of 'Fourier imaging'.

Two-dimensional considerations

The derivation of (1.181) in ¶1.7.1 and ¶1.7.2 is for the two dimensional case. However, the derivation may be easily extended to the three-dimensional case, as shown by Wolf(1969) for projection measurements and Morse and Ingard (1968 ¶8.1) for far-field measurements. The components of the position vector $\mathbf{r}$ for an arbitrary point are denoted by x and y for the coordinate system of (1.8). It is convenient to introduce a second coordinate system with unit vectors $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\eta}}$ which is rotated with respect to the first in such a way that

$$\hat{\mathbf{k}} = \hat{\boldsymbol{\eta}} \, . \tag{1.182}$$

The coordinates of P are (ξ,η) in the rotated system. The relationship of the two systems is illustrated in Fig. 1.5.

For two-dimensional scattering, the Green's function for an outgoing wave in the reference medium (¶1.3.5) is (Morse and Feshbach 1953 p.891)

$$G_o(\mathbf{r}|\mathbf{r}') = -(j/4)\, H_o^{(2)}(k|\mathbf{r} - \mathbf{r}'|) \tag{1.183}$$

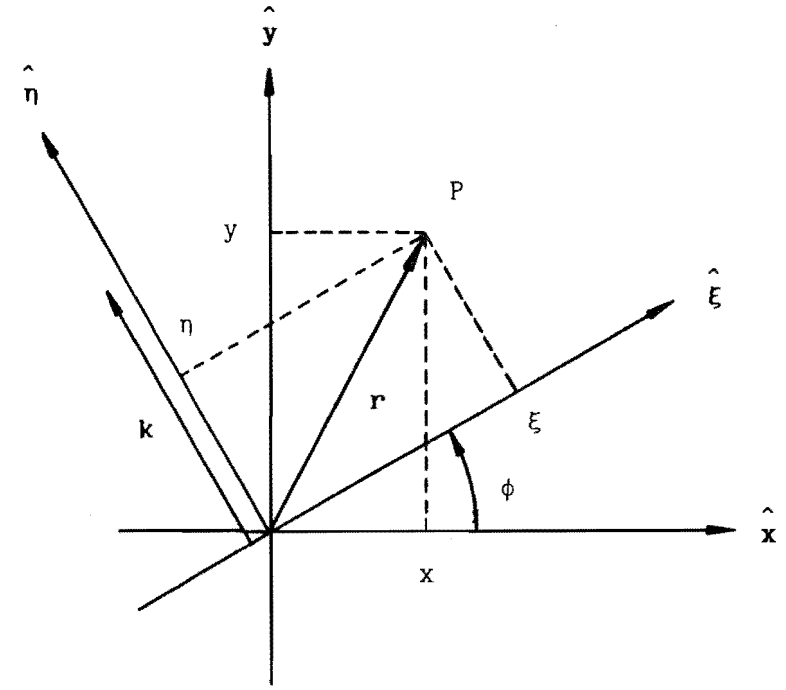


Fig. 1.5 Two-dimensional scattering geometry. The point P may be referred to by its position vector **r** or its coordinates in one of two coordinate systems. In the fixed system, defined by the unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, the coordinates are (x,y). The rotated system, having $\hat{\xi}$ and $\hat{\eta}$ as unit vectors, is aligned so that $\hat{\eta}$ is identical to the unit propagation vector $\hat{\mathbf{k}}$ of the incident wave. The angle between $\hat{\xi}$ and $\hat{\mathbf{x}}$ is denoted by ϕ. In this system P has coordinates (ξ,η). After Lewitt and Bates (1978).

where $H_0^{(2)}$ is the zeroth order Hankel function of the second kind.

1.7.1 Projection Measurements

In this class of measurements, ψ_{obj} is obtained on a line ℓ which is parallel to the ξ axis and intersects the η axis at $\eta = R$. For measurements in $T_{meas}^{(PF)}$, R is positive while for measurements in $T_{meas}^{(PB)}$

R is negative. The geometry for projection measurements in object space is shown in Fig. 1.6.

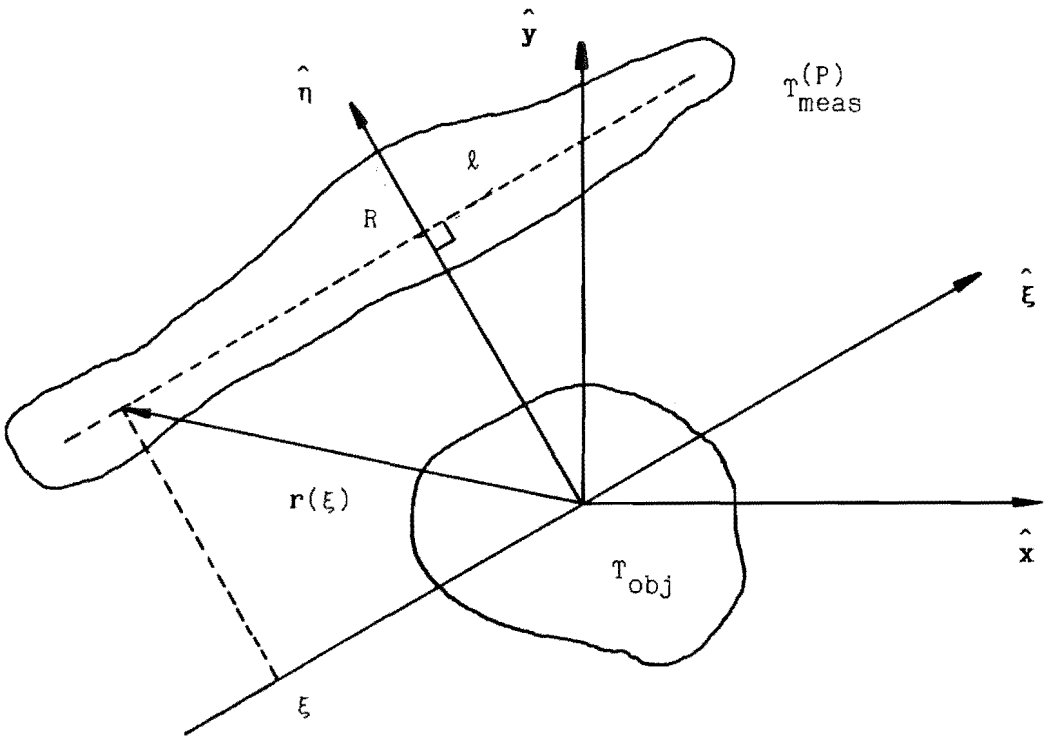


Fig. 1.6 The geometry for projection measurements. Measurements of ψ_{obj} are made on the line ℓ which intersects with the $\hat{\eta}$ axis at R. R may be positive or negative. A point on the line with coordinate ξ is denoted by $\mathbf{r}(\xi)$.

The Hankel function $H_0^{(2)}$ has the integral representation (Morse and Feshbach 1953 p.823; Abramowitz and Stegun 1965 p.360)

$$H_0^{(2)}(k|\mathbf{r}|) = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\tilde{\xi}}{\tilde{\eta}(\tilde{\xi})}\exp[-jK(\tilde{\xi})\cdot\mathbf{r}] \qquad (1.184)$$

where

$$\tilde{\eta}(\tilde{\xi}) = [k^2/4\pi^2 - \tilde{\xi}^2]^{1/2}\ , \qquad (1.185)$$

$$\mathbf{K}(\tilde{\xi}) = 2\pi[\tilde{\xi}\hat{\xi} \pm \tilde{\eta}(\tilde{\xi})\hat{\eta}] \ , \tag{1.186}$$

and

$$\pm \mathbf{r}\cdot\hat{\eta} > 0 \ . \tag{1.187}$$

In (1.186), (1.187), and throughout this section the upper sign corresponds to the case $T^{(P)}_{meas} = T^{(PF)}_{meas}$ while the lower sign corresponds to the case $T^{(P)}_{meas} = T^{(PB)}_{meas}$.

Physically, the integral representation (1.184) is an expansion of the Hankel function as a sum of plane waves, each characterised by the wave vector $\mathbf{K}(\tilde{\xi})$. When $2\pi\tilde{\xi}$ lies outside the interval $[-k,k]$ the wave vector is complex and the waves are evanescent, decaying exponentially in amplitude with distance. For $2\pi\tilde{\xi}$ within the interval $[-k,k]$, $\mathbf{K}(\tilde{\xi})$ is real and lies on a semicircle in wave vector space as shown in Fig. 1.7.

Upon substituting (1.183) into (1.163) and using (1.164), the first order scattering within the Born approximation is found to be given by

$$\psi_1^{(B)}(\mathbf{r},\omega) = -\frac{j}{4}\int_{T_{obj}} q_{obj}(\mathbf{r}')\ \psi_{gen}(\mathbf{r}',\omega)\ H_o^{(2)}(k|r-r'|)\ d\mathbf{r}' \ . \tag{1.188}$$

When the integral representation (1.184) is substituted into (1.188) a two-dimensional spatial Fourier transform may be identified in the right hand side,

$$\begin{aligned}\psi_1^{(B)}(\mathbf{r},\omega) &= \frac{j}{4\pi} A_{gen}(\omega) \int_{-\infty}^{\infty} \frac{d\tilde{\xi}}{\tilde{\eta}(\tilde{\xi})} \int_{T_{obj}} d\mathbf{r}'\ q_{obj}(\mathbf{r}') \\ &\qquad \times \exp[-j(\mathbf{k} - \mathbf{K}(\tilde{\xi}))\cdot\mathbf{r}']\ \exp[-j\mathbf{K}(\tilde{\xi})\cdot\mathbf{r}] \\ &= \frac{j}{4\pi} A_{gen}(\omega) \int_{-\infty}^{\infty} \frac{d\tilde{\xi}}{\tilde{\eta}(\tilde{\xi})}\ \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi})}{2\pi}\right) \exp[-j\mathbf{K}(\tilde{\xi})\cdot\mathbf{r}]\end{aligned} \tag{1.189}$$

where $\tilde{q}_{obj}$ is the aforementioned Fourier transform of q_{obj}. The integral on the right hand side of (1.189) may be inverted by taking the one-dimensional inverse Fourier transform of $\psi_1^{(B)}$ along the line ℓ. The result, evaluated at spatial frequency $\tilde{\xi}'$, is

$$\int_{-\infty}^{\infty} \psi_1^{(B)}(\mathbf{r},\omega)\ \exp[j2\pi\tilde{\xi}'\xi]\ d\xi$$

$$= \int_{-\infty}^{\infty} \psi_1^{(B)}(\mathbf{r},\omega)\ \exp[j\mathbf{K}'(\tilde{\xi}')\cdot\mathbf{r}(\xi)]\ d\xi$$

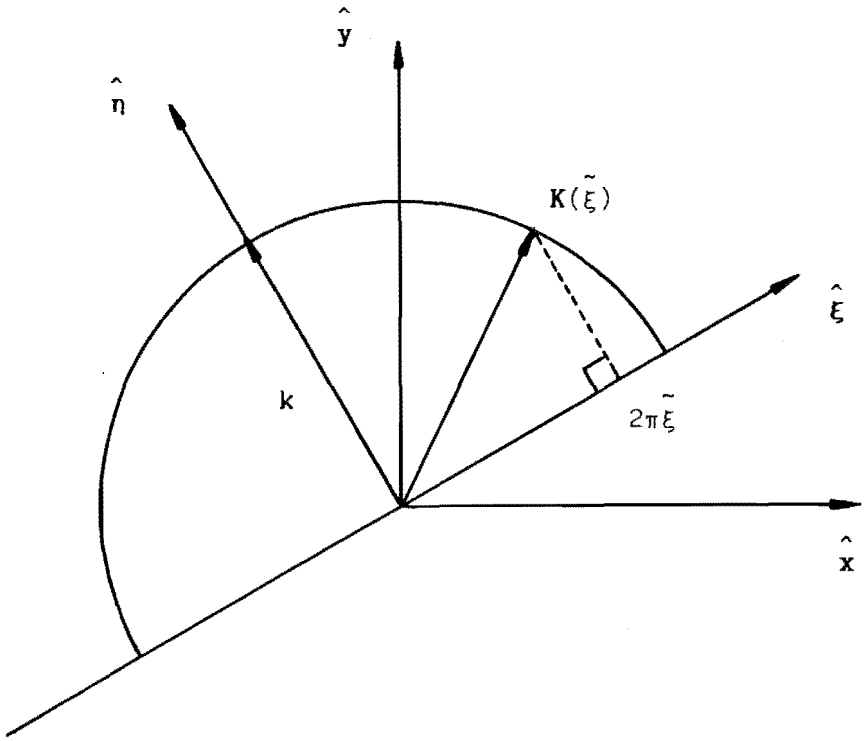


Fig. 1.7 The plane wave expansion. The Green's function can be expressed as a series of plane waves of wave vector $\mathbf{K}(\tilde{\xi})$. For $2\pi\tilde{\xi}$ in the interval $[-k,k]$, $\mathbf{K}(\tilde{\xi})$ is real and lies on the semicircle shown.

$$= \frac{j}{4\pi} A_{gen}(\omega) \int_{-\infty}^{\infty} \frac{d\tilde{\xi}}{\tilde{\eta}(\tilde{\xi})} \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi})}{2\pi}\right)$$

$$\times \int_{-\infty}^{\infty} \exp\left[-j\left(\mathbf{K}(\tilde{\xi}) - \mathbf{K}'(\tilde{\xi}')\right)\cdot\mathbf{r}(\xi)\right] d\xi$$

$$= \frac{j}{4\pi} A_{gen}(\omega)\, \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi}')}{2\pi}\right)$$

$$\times \frac{\exp\left[-\pm jK(\tilde{\xi}')\cdot\mathbf{r}(\xi=0)\right]}{\tilde{\eta}(\tilde{\xi}')} \qquad (1.190)$$

where the vector $\mathbf{K}'(\tilde{\xi}')$ is defined by

$$\mathbf{K}'(\tilde{\xi}') = 2\pi\tilde{\xi}'\hat{\xi} \tag{1.191}$$

and $\mathbf{r}(\xi)$ is the point on the line ℓ given by

$$\mathbf{r}(\xi) = \xi\hat{\xi} + R\hat{\eta} \ . \tag{1.192}$$

In deriving (1.190), the formula

$$\begin{aligned}
&\int_{-\infty}^{\infty} \exp[-j(\mathbf{K}(\tilde{\xi}) - \mathbf{K}'(\tilde{\xi}'))\cdot\mathbf{r}(\xi)]\, d\xi \\
&= \int_{-\infty}^{\infty} \exp[-j2\pi((\tilde{\xi} - \tilde{\xi}')\xi \pm \tilde{\eta}(\tilde{\xi})\eta)]\, d\xi \\
&= \delta(\tilde{\xi} - \tilde{\xi}')\, \exp[-\pm j2\pi\tilde{\eta}(\tilde{\xi})\eta] \\
&= \delta(\tilde{\xi} - \tilde{\xi}')\, \exp[-\pm j\mathbf{K}(\tilde{\xi})\cdot\mathbf{r}(\xi=0)]
\end{aligned} \tag{1.193}$$

is used along with the sifting property (1.127) of the δ function.

A similar relationship to (1.190) may be derived within the Rytov approximation. To do this, the integrand of the right hand side of (1.190) is replaced by the complex phase difference between the total wave and the incident wave,

$$\Delta\phi^{(R)} = \phi^{(R)} - \phi_{gen} \ , \tag{1.194}$$

and the left hand side is reevaluated. This is facilitated by noting that the phase difference is very simply related to the first order scattering in the Born approximation. Invoking (1.173), this relation is given by

$$\Delta\phi^{(R)} = \frac{\psi_1^{(B)}}{\psi_{gen}} \ . \tag{1.195}$$

The one-dimensional inverse Fourier transform of $\Delta\phi^{(R)}$ along ℓ is

$$\begin{aligned}
&\int_{-\infty}^{\infty} \Delta\phi^{(R)}(\mathbf{r}(\xi),\omega)\, \exp[j\mathbf{K}'(\tilde{\xi}')\cdot\mathbf{r}(\xi)]\, d\xi \\
&= \int_{-\infty}^{\infty} \frac{\psi_1^{(B)}}{A_{gen}} \exp[j\mathbf{k}\cdot\mathbf{r}(\xi)]\, \exp[j\mathbf{K}'(\tilde{\xi}')\cdot\mathbf{r}(\xi)]\, d\xi
\end{aligned}$$

$$= \int_{-\infty}^{\infty} \frac{\psi_1^{(B)}}{A_{gen}} \exp[j\mathbf{k}\cdot\mathbf{r}(\xi=0)] \exp[j\mathbf{K}'(\tilde{\xi}')\cdot\mathbf{r}(\xi)] \, d\xi$$

$$= \frac{j}{4\pi} \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi}')}{2\pi}\right) \frac{\exp[-j(\pm\mathbf{K}(\tilde{\xi}') - \mathbf{k})\cdot\mathbf{r}(\xi=0)]}{\tilde{\eta}(\tilde{\xi}')} \qquad (1.196)$$

where (1.195) and (1.190) are invoked and it is noted that ℓ is perpendicular to the wave vector $\mathbf{k}$.

(1.196) resembles (1.190) and this similarity may be exploited to condense the two equations into a single formula. To this end, scaled versions of the data available on the line ℓ are defined by

$$\tilde{\chi}^{(P)}(\xi,\omega) = \begin{cases} \psi_1^{(B)}[\mathbf{r}(\xi),\omega] & \text{Born} \quad (1.197) \\ \Delta\phi^{(R)}[\mathbf{r}(\xi),\omega] \exp[-j\mathbf{k}\cdot\mathbf{r}(\xi=0)] & \text{Rytov} \quad (1.198) \end{cases}$$

and

$$\tilde{\chi}^{(P)}(\xi,\tilde{\xi},\omega) = -j4\pi\tilde{\eta}(\tilde{\xi}) \exp[\pm j\mathbf{K}(\tilde{\xi})\cdot\mathbf{r}(\xi=0)] \; \tilde{\chi}^{(P)}(\xi,\omega) \; . \qquad (1.199)$$

The data $\tilde{\chi}^{(P)}(\xi,\omega)$ and $\tilde{\chi}^{(P)}(\xi,\omega,\tilde{\xi})$ are scaled projections of the plane wave through the object onto the line ℓ. With the aid of definitions (1.197), (1.198), and (1.199), (1.190) and (1.196) are represented by the single formula

$$\int_{-\infty}^{\infty} \tilde{\chi}^{(P)}(\xi,\tilde{\xi},\omega) \exp[j2\pi\tilde{\xi}\xi] \, d\xi = \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi}')}{2\pi}\right) \qquad (1.200)$$

The theorem (1.200) states that the one-dimensional inverse Fourier transform of the projection data $\tilde{\chi}^{(P)}$ is equal to a slice through the two-dimensional Fourier transform of the object on the semicircle $\mathbf{k} - \mathbf{K}(\tilde{\xi})$ shown in Fig. 1.8. Denoting the inverse Fourier transform of $\tilde{\chi}^{(P)}(\xi,\tilde{\xi},\omega)$ with respect to ξ by $\chi^{(P)}(\tilde{\xi},\omega)$, (1.200) may be written

$$\chi^{(P)}(\tilde{\xi},\omega) = \tilde{q}_{obj}\left(\frac{\mathbf{k} - \mathbf{K}(\tilde{\xi}')}{2\pi}\right) \; . \qquad (1.201)$$

Further, if the vector $\mathbf{u}$ is defined by

$$\mathbf{u}(\tilde{\xi},\omega) = [\mathbf{k} - \mathbf{K}(\tilde{\xi})]/2\pi \; , \qquad (1.202)$$

(1.201) can be written in the form (1.181) of the generalised projection theorem.

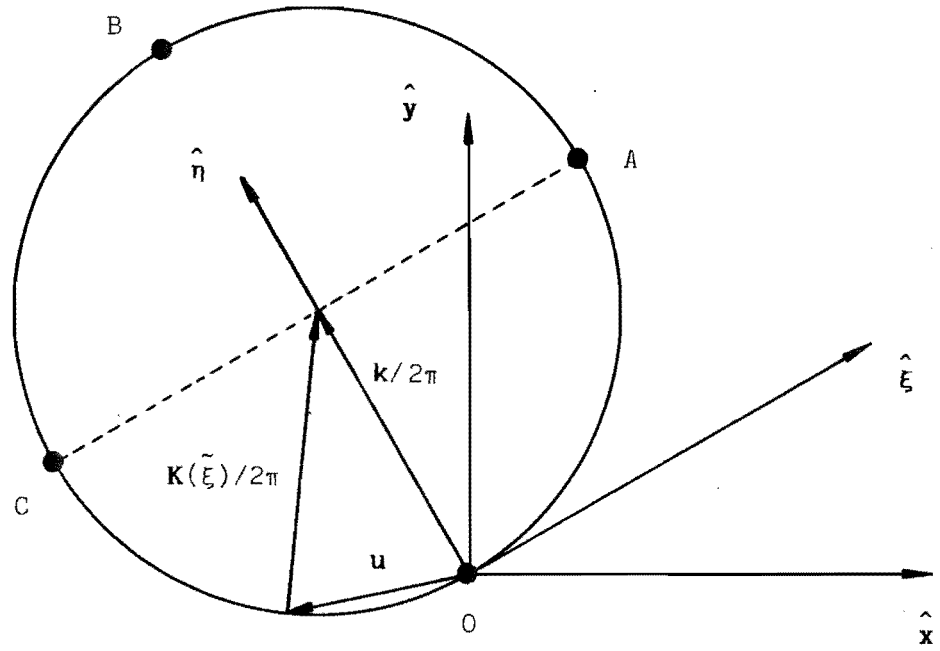


Fig. 1.8 The geometry of Fourier space for projection measurements. The two-dimensional spatial Fourier transform of q_{obj} is obtained for points on the semicircle COA or ABC, for measurements in $\psi^{(PF)}_{meas}$ and $\psi^{(PB)}_{meas}$ respectively. As $2\pi\tilde{\xi}$ takes upon the values $-k$ to k, $\mathbf{K}(\tilde{\xi})$ rotates and the position vector $\mathbf{u}$ sweeps out the appropriate semicircle.

The solution of the remote probing problem based upon (1.201) is called 'diffraction tomography' (Mueller 1980). The special case of diffraction tomography, which results when diffraction effects are ignored, is called 'computed tomography (CT)' (Bates et al. 1983).

A result similar to (1.200) is derived by Wolf (1969) for the three-dimensional case within the Born approximation. The Rytov approximation is considered in Iwata and Nagata (1975). Devaney (1982) gives a derivation similar to that presented here, but the formulas obtained within the Rytov approximation are incorrect with respect to the amplitude of the incident wave and the case $R < 0$ is not considered. The

result (1.200) for $T^{(PF)}_{meas}$ is referred to as the 'Fourier diffraction projection theorem' by Pan and Kak (1983). Mueller et al. (1979) review diffraction tomography and discuss the validity of the Born and Rytov approximations. A recent review of the subject is provided by Slaney and Kak (1983). Devaney and Beylkin (1984) provide reconstruction algorithms for diffraction tomography within the Born approximation for plane and cylindrical insonification. These cases are analogous to parallel and fan beam reconstructions in CT (Garden 1984 ¶4.2). Lu et al. (1984) discuss the use of diffraction tomography to obtain reconstructions which are averaged along one coordinate.

CT may be obtained from diffraction tomography by considering the high frequency limit. The semicircle described by (1.186) becomes a straight line,

$$\mathbf{K}(\tilde{\xi}) \to 2\pi\tilde{\xi}\hat{\tilde{\xi}} \pm \mathbf{k} \quad \text{as } k \to \infty \,, \tag{1.203}$$

and the data (1.197), (1.198) are calculated from

$$\tilde{\chi}^{(P)}(\tilde{\xi},\omega) = -j2k \begin{cases} \exp[jkR]\ \psi_1^{(B)}(\mathbf{r}(\xi),\omega) & \text{Born} \quad (1.204) \\ \Delta\phi^{(B)}(\mathbf{r}(\xi),\omega) \,. & \text{Rytov} \quad (1.205) \end{cases}$$

The data within the Rytov approximation at high frequencies is simply related to the refractive index. A first order asymptotic expansion for the total wave is (Born and Wolf 1964 ¶3.1.3)

$$\psi(\mathbf{r}) = A_{gen}\,[n(\mathbf{r})]^{-1/2} \exp[-jk\int^{\mathbf{r}} n(\mathbf{r}')\, d\ell] \tag{1.206}$$

where the integral is along a ray path. The complex phase difference defined by (1.194) is given by

$$\Delta\phi^{(R)}(\mathbf{r}) = -\tfrac{1}{2}\ln(n) + jk\int^{\mathbf{r}} [1 - n(\mathbf{r}')]\, d\ell \,. \tag{1.207}$$

At high enough frequencies the second term will dominate the first. The data $\Delta\phi^{(R)}$ are a projection of the refractive index. When attenuation is considered, the refractive index is complex and projections of the absorbitivity of the object are obtained. For X-ray CT the refractive index is effectively unity and so only projections of absorbitivity are considered and these are along straight ray paths. The remote probing problem for ultrasound is considerably more complicated than for X-rays because there is an additional unknown - the refractive index and the projections are generated along curved ray paths.

1.7.2 Far-Field Measurements

A measurement in the far-field measurement class is obtained at an observation point $\mathbf{r}$ in $T_{meas}^{(F)}$ as shown in Fig. 1.9. A typical source point within the object T_{obj} is denoted by $\mathbf{r}'$. At large distances from the source point the Green's function (1.183) has a simple exponential form. For large values of the argument, the Hankel function $H_0^{(2)}$ is given approximately by (Morse and Feshbach 1953 p.624)

$$H_0^{(2)}(kr) = \left[\frac{2}{\pi kr}\right]^{1/2} \exp[-j(kr - \pi/4)] \quad \text{as } kr \to \infty \ . \tag{1.208}$$

The distance $|\mathbf{r} - \mathbf{r}'|$ between the source point $\mathbf{r}'$ and the observation point $\mathbf{r}$ can be approximated by truncating the binomial series

$$(1+x)^{1/2} = 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \ldots \ . \tag{1.209}$$

The result is

$$\begin{aligned} |\mathbf{r} - \mathbf{r}'| &= \left[(\mathbf{r} - \mathbf{r}')\cdot(\mathbf{r} - \mathbf{r}')\right]^{1/2} \\ &= r\left[1 - 2\mathbf{r}\cdot\mathbf{r}'/r^2 + \mathbf{r}'\cdot\mathbf{r}'/r^2\right]^{1/2} \\ &\approx r\left[1 - \hat{\mathbf{r}}\cdot\left(\frac{\mathbf{r}'}{r}\right) + \frac{r'^2}{2r^2} - \frac{(\mathbf{r}\cdot\mathbf{r}')^2}{2r^4}\right] \end{aligned} \tag{1.210}$$

where r, r', and $\hat{\mathbf{r}}$ are defined by

$$r = |\mathbf{r}| \ , \tag{1.211}$$

$$r' = |\mathbf{r}'| \ , \tag{1.212}$$

and

$$\hat{\mathbf{r}} = \mathbf{r}/r \ . \tag{1.213}$$

Terms of higher order than $(\mathbf{r}\cdot\mathbf{r}')^2/2r^4$ are neglected in (1.210). The region in which terms of order $(r'/r)^2$ have a negligible effect on $G_0(\mathbf{r}|\mathbf{r})$ is known as the 'Fraunhofer' region or the 'far-field'. The usual condition is that the phase of the complex exponential in (1.208) is not altered by more than $\pi/8$ radians (Blake 1966 p.122),

$$\left|k\left(\frac{r'^2}{2r} - \frac{\mathbf{r}\cdot\mathbf{r}'}{2r^3}\right)\right| \leq \frac{\pi}{8} \ , \tag{1.214}$$

or

$$\left|k\frac{r'^2}{2r}\left(1 - \cos^2(\theta)\right)\right| \leq \frac{\pi}{8} \tag{1.215}$$

where θ is the angle between $\mathbf{r}$ and $\mathbf{r}'$. For arbitrary directions, (1.215)

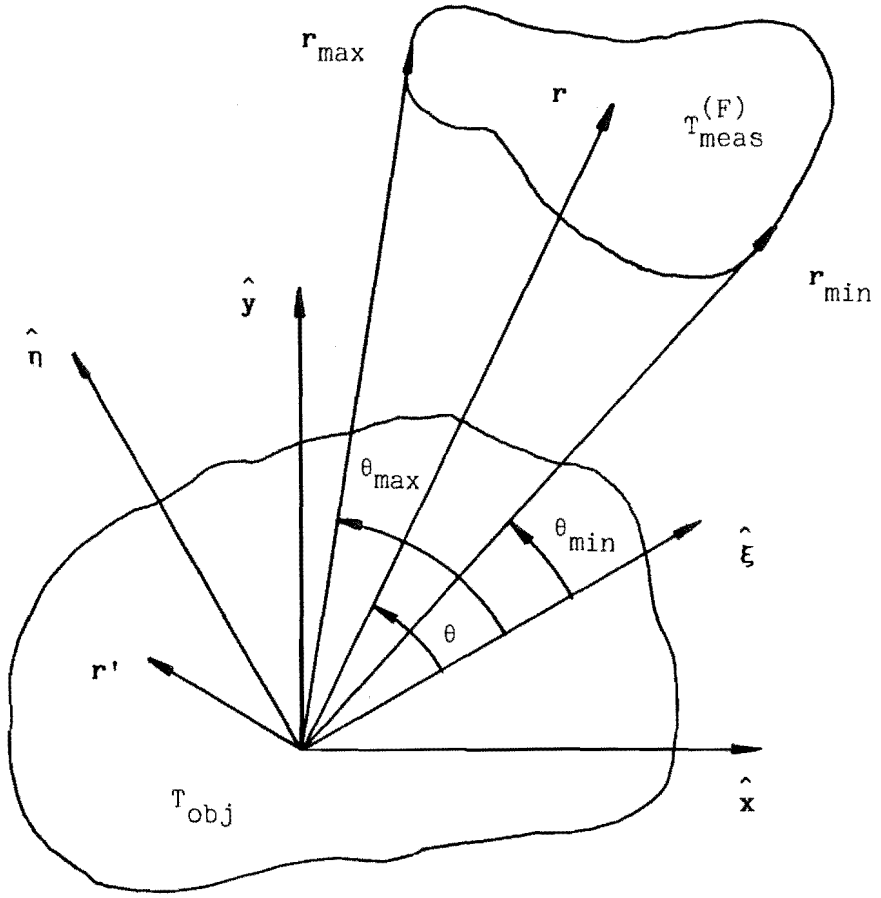


Fig. 1.9 The geometry of far-field measurements. Scattering from the source point $\mathbf{r}'$ within the object is measured at the observation point $\mathbf{r}$ in $T^{(F)}_{meas}$. The angle θ at which scattering at $\mathbf{r}$ is observed is measured with respect to $\hat{\xi}$. Scattering can be measured in a range of directions from θ_{min} to θ_{max}. The corresponding points at which the end points of this range are achieved are denoted by $\mathbf{r}_{min}$ and $\mathbf{r}_{max}$.

becomes

$$\frac{r}{r'} \geq \frac{8r'}{\lambda} \quad . \tag{1.216}$$

When the truncated version of (1.210) is substituted into (1.208), and the approximation $|\mathbf{r} - \mathbf{r}'| \approx r$ is made in the amplitude factor, the Green's function is found to be approximately

$$G_0(\mathbf{r}|\mathbf{r}') \approx G_0(\mathbf{r}) \exp[jk\hat{\mathbf{r}}\cdot\mathbf{r}'] \tag{1.217}$$

where

$$G_0(\mathbf{r}) = -\frac{j}{4}\left[\frac{2}{\pi kr}\right]^{1/2} \exp[j\tfrac{\pi}{4}] \exp[-jkr] \quad . \tag{1.218}$$

When the expressions for the incident plane wave (1.180) and the approximate Green's function (1.217) are substituted into (1.163), the first order scattering in the Born approximation is found to be

$$\begin{aligned}\psi_1^{(B)}(\mathbf{r}) &= \int_{T_{obj}} q_{obj}(\mathbf{r}')\, A_{gen} \exp[-j\mathbf{k}\cdot\mathbf{r}']\, G_0(\mathbf{r}) \exp[jk\hat{\mathbf{r}}\cdot\mathbf{r}']\, d\mathbf{r}' \\ &= A_{gen}\, G_0(\mathbf{r})\, \tilde{q}_{obj}\left(\frac{\mathbf{k} - k\hat{\mathbf{r}}}{2\pi}\right) \\ &= A_{gen}\, G_0(\mathbf{r})\, \tilde{q}_{obj}(\mathbf{u})\end{aligned} \tag{1.219}$$

where

$$\mathbf{u} = \frac{k}{2\pi}(\hat{\mathbf{k}} - \hat{\mathbf{r}}) \tag{1.220}$$

and $\tilde{q}_{obj}$ is the two-dimensional spatial Fourier transform of q_{obj}.

A similar result may be obtained with the Rytov approximation. The relationship (1.219) is substituted into (1.195), to obtain an expression for the complex phase difference of the wave in the far-field,

$$\Delta\phi^{(R)}(\mathbf{r}) = G_0(\mathbf{r}) \exp[j\mathbf{k}\cdot\mathbf{r}]\, \tilde{q}_{obj}(\mathbf{u}) \tag{1.221}$$

The results (1.221) and (1.219) may be combined by a method analogous to that employed in ¶1.7.1. The data χ obtained in the far-field is defined by

$$\chi^{(F)}(\hat{\mathbf{r}},\omega) = \begin{cases} \dfrac{\psi^{(B)}{}_1(\mathbf{r},\omega)}{A_{gen}(\omega)\ G_0(\mathbf{r},\omega)} & \text{Born} \quad (1.222) \\ \dfrac{\Delta\phi^{(R)}(\mathbf{r},\omega)}{\exp[j\mathbf{k}\cdot\mathbf{r}]\ G_0(\mathbf{r},\omega)} & \text{Rytov} \quad (1.223) \end{cases}$$

where the frequency dependence of the various factors has been shown explicitly. (1.219) and (1.221) can now both be represented by

$$\chi^{(F)}(\hat{\mathbf{r}},\omega) = \tilde{q}_{obj}\left[\frac{k}{2\pi}(\hat{\mathbf{k}} - \hat{\mathbf{r}})\right] . \qquad (1.224)$$

With $\mathbf{u}$ defined by (1.220), (1.224) has the form of the generalised projection theorem (1.181).

1.7.3 A Comparison of Far-Field and Projection Measurements

The equations relating the data to the object for far-field and projection measurements are very similar. Both relate the data to a slice through the object's spatial Fourier transform.

To facilitate an accurate reconstruction of the object, its Fourier transform needs to be known over as large a region as possible. This region is called the 'Fourier space coverage' for the object. The parameters determining the slice are the direction and frequency of the incident wave and a third parameter which is either the direction of observation for a far-field measurement or a spatial frequency for a projection measurement. Since Fourier space is two-dimensional at least two of the parameters must be varied if Fourier space is to be sampled adequately.

In Fourier imaging when the data is obtained for all directions at a fixed frequency the slice is a complete circle. This circle is a section through what is known as the 'Ewald sphere' (Guinier 1963 ¶1). The radius of the circle is $k/2\pi$ and the centre is at $\mathbf{k}/2\pi$ so the circle intersects with the origin as shown in Fig. 1.10. When measurements are made for a range of directions from $\hat{\mathbf{r}}_{min}$ to $\hat{\mathbf{r}}_{max}$, as shown in Fig. 1.10, data is available only on the arc AB in Fourier space shown in Fig. 1.11. The coverage of Fourier space may be extended by making measurements over a range of wave numbers from k_{min} to k_{max}. The extended coverage is depicted as a shaded region in Fig. 1.12. This region is bounded by arcs on one opposing pair of sides and by straight lines on the remaining pair of sides. That these sides are in fact straight lines can be seen from (1.220) when k is varied with $\hat{\mathbf{r}}$ equal to $\hat{\mathbf{r}}_{min}$ or $\hat{\mathbf{r}}_{max}$. In general, the 'measurement aperture' is defined to be the range of frequencies and

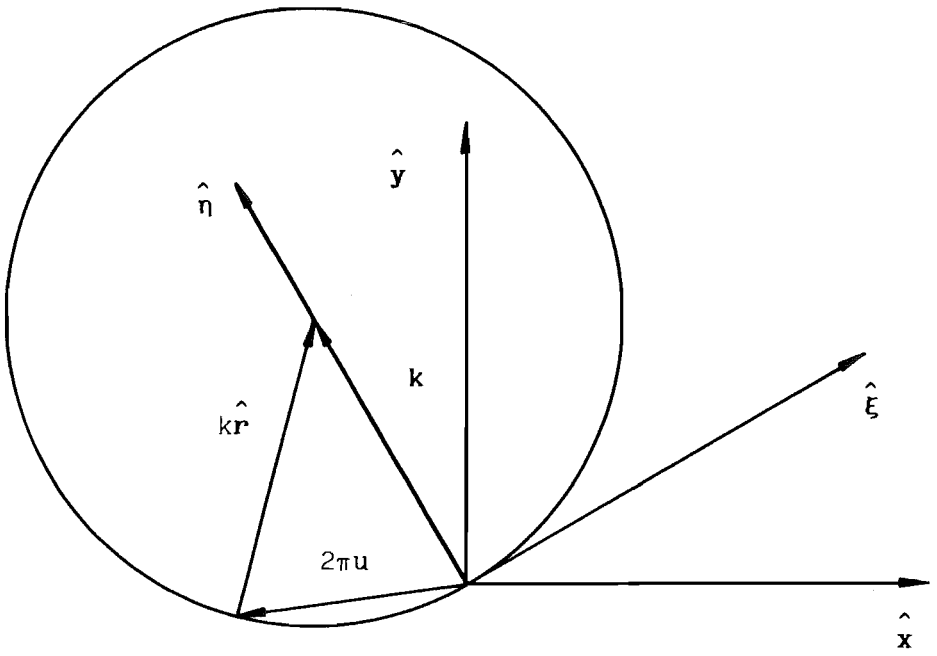


Fig. 1.10 Wave vector space for far-field measurements. Spatial frequency information about the object is available for points on the circle, which is a cross-section of the Ewald sphere. The circle is generated as the unit vector pointing to the observation point $\hat{r}$ sweeps around all directions.

directions for which measurements are obtained. Fig. 1.13 is a pictorial tabulation of the Fourier space coverages corresponding to various measurement apertures.

The equation governing diffraction tomography (1.200) relates the data to the object's spatial Fourier transform evaluated on a semicircle. Comparison of (1.185) and (1.186) with (1.220) shows that the Fourier space coverage obtained from a single projection measurement is equivalent to the coverage obtained for a far-field measurement aperture ranging from $-\hat{\xi}$ to $\hat{\xi}$. This coverage is shown in Fig. 1.14.

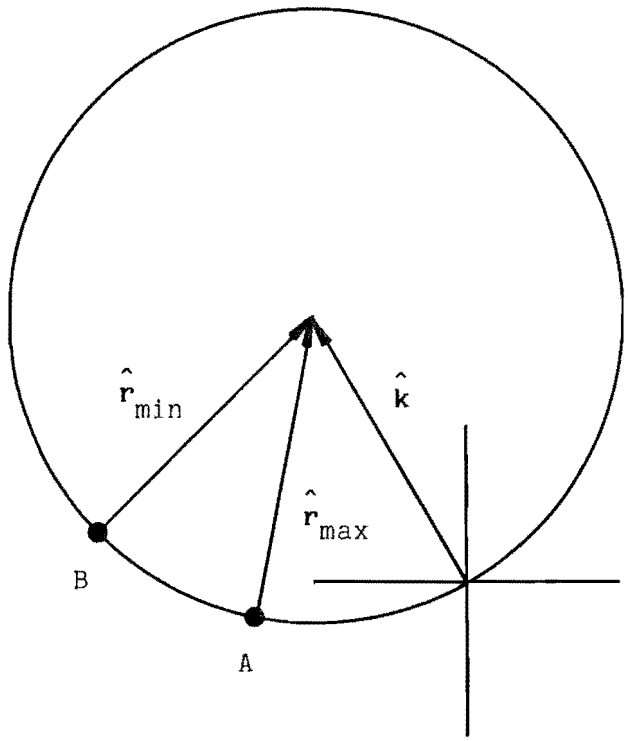


Fig. 1.11 Fourier space coverage for monochromatic measurements in a limited aperture. For the range of directions shown in Fig. 1.10, the Fourier transform of the object is obtained on the arc AB.

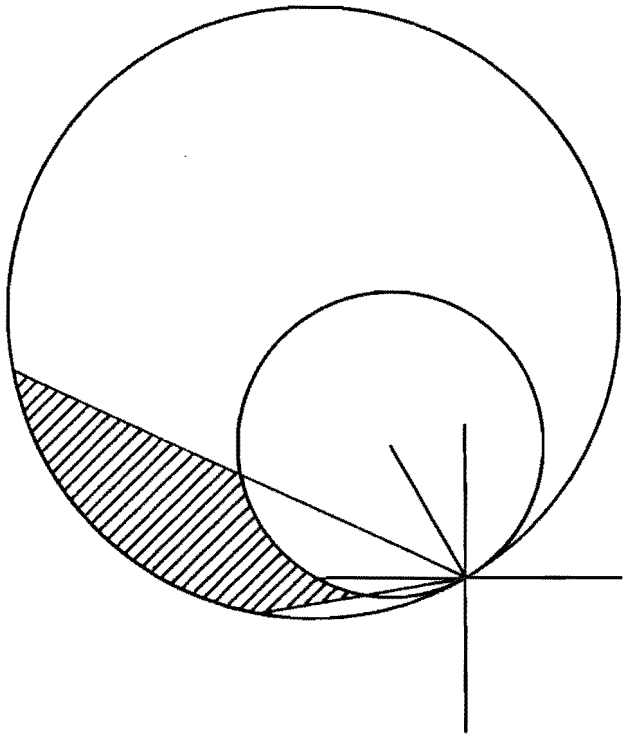


Fig. 1.12 Fourier space coverage for multi-frequency measurements in a limited aperture. The spatial Fourier transform of the object is obtained in the shaded region. The radius of the smaller and larger circles are $k_{min}/2\pi$, and $k_{max}/2\pi$ respectively.

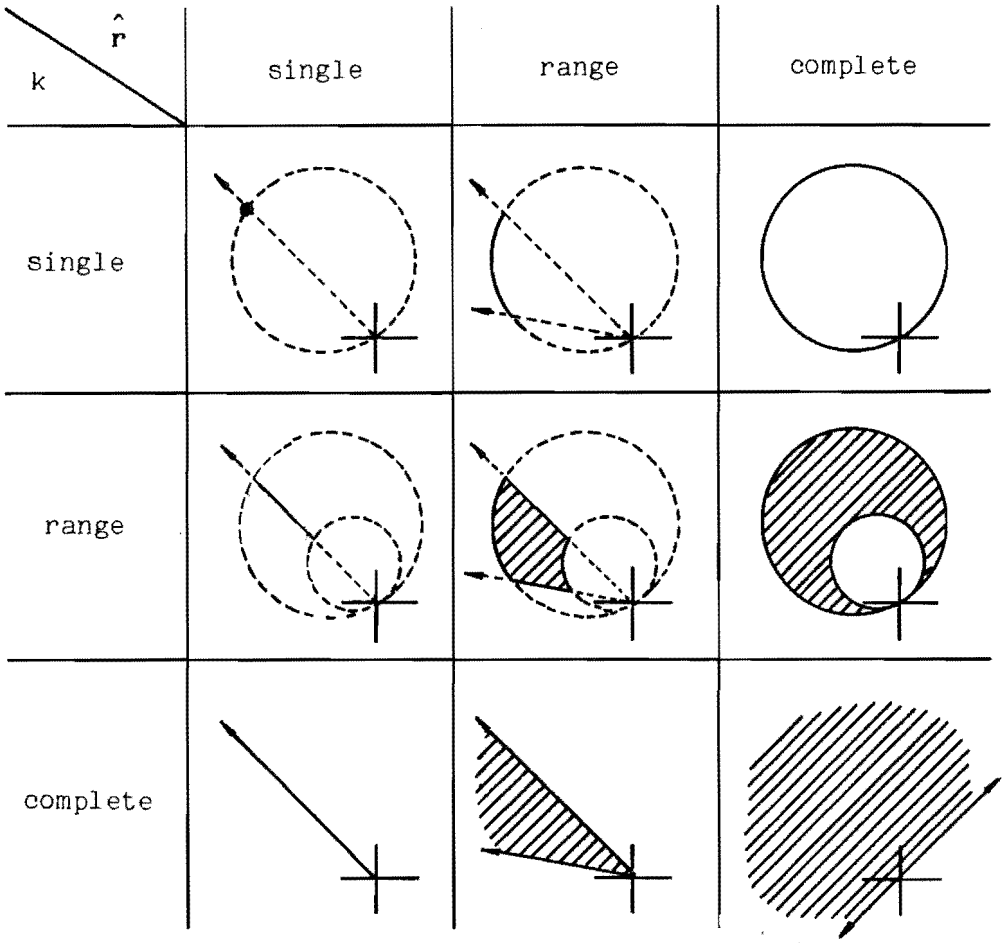


Fig. 1.13 Fourier space coverage for various measurement apertures. The columns correspond to measurements obtained for: a single direction, a range of directions and all directions respectively. Similarly, the rows correspond to measurements obtained: at a single frequency, at a range of frequencies and at all frequencies.

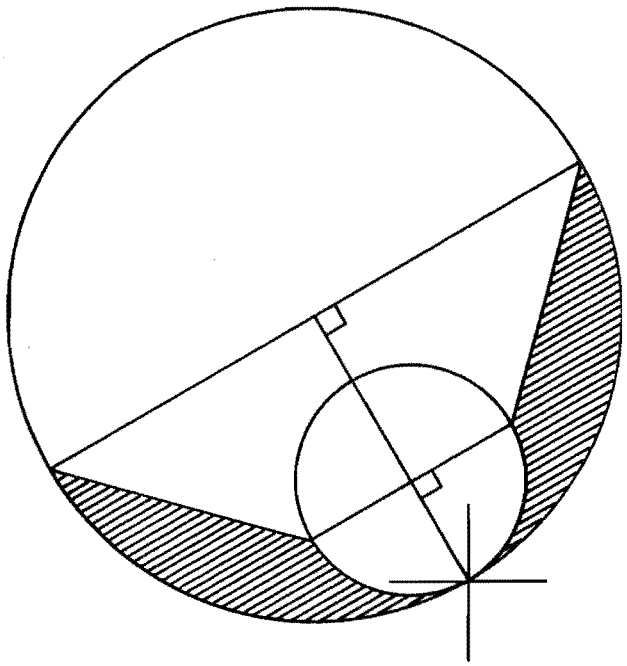


Fig. 1.14 Fourier space coverage for a projection. The shaded region is the coverage for a set of projection measurements obtained at a range of frequencies.

1.8 INVERSION OF THE FOURIER INTEGRAL

The Fourier integral in (1.201) and (1.224) may be inverted using the formula (1.58) for the inverse Fourier transform. The resulting equation indicates that in principle the object may be obtained from the data according to

$$q_{obj}(\mathbf{r}) = \int_{T_{\mathbf{u}}} \chi(\mathbf{u}) \exp[j2\pi \mathbf{u}\cdot\mathbf{r}]\, d\mathbf{u} \qquad (1.224)$$

where $d\mathbf{u}$ is a volume element in Fourier space and $\hat{\mathbf{r}}$ is related to $\mathbf{u}$ by (1.220). As is explained in ¶1.8.1 and ¶1.8.2, there are several reasons why (1.224) can only be computed approximately. Lewitt and Bates

(1978) discuss these difficulties with particular reference to X-ray CT.

For CT, the task of calculating $q_{obj}(\mathbf{r})$ from (1.224) is known as 'image reconstruction from projections' (Peters 1973, Lewitt 1977). Solution techniques for the remote probing problem in general are discussed in Bates et al. (1983) and Garden (1984). The particular problems posed by curved ray paths in ultrasonic CT are considered by McKinnon (1980).

1.8.1 Finite Fourier Space Coverage

One reason why q_{obj} can only be computed approximately from (1.224) is that the Fourier space coverage is inevitably incomplete. Even if the measurement aperture extends over all directions and the object is insonified from a number of different angles the range of frequencies employed is always limited by physical constraints. When the missing data are assumed to be zero, the estimate $\hat{q}_{obj}$ obtained from (1.224) is a filtered version of q_{obj}. To see this, let the Fourier space coverage be T_χ and define a function $\tilde{C}_\chi$ on Fourier space which is 1 in T_χ and 0 elsewhere,

$$\tilde{C}_\chi(\mathbf{u}) = \begin{cases} 1 & \text{if } \mathbf{u} \text{ is in } T_\chi \\ 0 & \text{if } \mathbf{u} \text{ is not in } T_\chi \end{cases} . \tag{1.225}$$

The estimate of the object that is actually calculated is

$$\hat{q}_{obj}(\mathbf{r}) = \int_{T_\mathbf{u}} \chi(\mathbf{u})\, \tilde{C}_\chi(\mathbf{u})\, \exp[j2\pi\mathbf{u}\cdot\mathbf{r}]\, d\mathbf{u} . \tag{1.226}$$

With aid of the convolution theorem (1.63) it can be seen that the estimate is a filtered version of the true object image,

$$\hat{q}_{obj} = q_{obj} \circledcirc C_\chi \tag{1.227}$$

where C_χ is the two-dimensional inverse Fourier transform of $\tilde{C}_\chi$. When (1.227) is compared with (1.144) it can be seen that Fourier imaging is a PSI image reconstruction technique (¶1.5.4). The effect of the finite Fourier space coverage is to blur the image with the PSF (¶1.5.4) C_χ.

1.8.2 Sampling Considerations

A second difficulty in computing q_{obj} from (1.224) arises when the numerical evaluation of (1.224) is considered. Fourier transforms are usually calculated using an algorithm called the 'Fast Fourier Transform (FFT)' (Brigham 1974). One of the constraints of this algorithm is that

images should be sampled regularly (¶1.2.5). The data obtained from projection or far-field measurements lie on a circular arc and so are not suitable for direct input to the FFT. An interpolation is necessary, if a two-dimensional FFT is to be used to implement (1.224).

To perform the interpolation some assumption has to be made about the nature of $\tilde{q}_{obj}$. If q_{obj} is of finite extent, $\tilde{q}_{obj}$ is band-limited (¶1.2.5) and q_{obj} may be reconstructed from its samples by sinc interpolation providing the samples satisfy the Nyquist theorem (1.73) and (1.74). However the sinc function is of infinite extent, so that the convolution can be computed only approximately by truncating the sinc function to a finite domain. Further truncation is usually necessary to reduce the computational effort. Therefore, in practice, the interpolation can only be performed approximately. An interpolation error at a particular point in Fourier space gives rise to an artefact which affects all of object space. The need to perform interpolation in Fourier space, and the consequent inaccuracies, are known as the interpolation difficulty. This difficulty is discussed by Bates (1981) in a general context, and by Heffernan and Bates (1982) for X-ray CT.

1.8.3 A Parametric Approach

In this subsection a reconstruction procedure is derived under the constraints that the direction of incidence is fixed and the observation direction and frequency are varied. Nahamoo et al. (1984) and Nahamoo and Kak (1982) employ a parametric approach to obtain an algorithm for diffraction tomography which requires no interpolation.

An alternative procedure to the two-dimensional FFT inversion of (1.224) can be derived by transforming the integral from $T_{\mathbf{u}}$ space to the parametric space (θ,k) where θ is the angle between the observation direction and the ξ axis (Fig. 1.10). In terms of θ the observation direction is

$$\hat{\mathbf{r}} = \cos(\theta)\hat{\boldsymbol{\xi}} + \sin(\theta)\hat{\boldsymbol{\eta}} \,. \tag{1.228}$$

Upon substituting (1.228) and (1.182) into (1.220), $\mathbf{u}$ is found to be given by

$$\mathbf{u} = u_{\xi}\hat{\boldsymbol{\xi}} + u_{\eta}\hat{\boldsymbol{\eta}} \tag{1.229}$$

where

$$u_{\xi} = -(k/2\pi)\cos(\theta) \tag{1.230}$$

and

$$u_\eta = (k/2\pi)\{1 - \sin(\theta)\} \ . \tag{1.231}$$

The area element $d\theta dk$ for the transformed integral is related to the area element $d\mathbf{u}$ by the Jacobian $\partial(u_\xi,u_\eta)/\partial(\theta,k)$ of the transformation from (u_ξ,u_η) to (θ,k) (Courant 1936 p.252)

$$d\mathbf{u} = du_\xi du_\eta = \frac{\partial(u_\xi,u_\eta)}{\partial(\theta,k)}\, d\theta dk \ . \tag{1.232}$$

The Jacobian is defined by (Courant 1936 p.143)

$$\frac{\partial(u_\xi,u_\eta)}{\partial(\theta,k)} = \frac{\partial u_\xi}{\partial\theta}\frac{\partial u_\eta}{\partial k} - \frac{\partial u_\xi}{\partial k}\frac{\partial u_\eta}{\partial\theta} \ . \tag{1.233}$$

For the transformation (1.230) and (1.231), it is given by

$$\frac{\partial(u_\xi,u_\eta)}{\partial(\theta,k)} = \frac{k}{4\pi^2}\{1 - \sin(\theta)\} \ . \tag{1.234}$$

The area element $d\theta dk$ is shown in Fig. 1.15. As θ increases from $-\pi/2$ to $\pi/2$ the Jacobian decreases to zero. This is to be expected since $T_{\mathbf{u}}$ is sampled more densely by the new parameters (θ,k) near the origin and so the data needs to be given less weight in the integral (1.224).

The integral (1.226) is transformed, using (1.229), (1.230), (1.231), (1.232), and (1.234), to

$$\hat{q}_{obj}(\mathbf{r}) = \int_{k_{min}}^{k_{max}} dk \int_{\theta_{min}}^{\theta_{max}} d\theta\, \frac{k}{4\pi^2}\{1 - \sin(\theta)\}\, \chi(k,\theta) \times \exp\{jk[-\xi\cos(\theta) + \eta\{1 - \sin(\theta)\}]\} \ . \tag{1.235}$$

1.8.4 Back-Projection Interpretation

The inversion of (1.224) can be pictured physically. The inverse Fourier kernel $\exp[j2\pi\mathbf{u}\cdot\mathbf{r}]$ represents a plane wave travelling in the direction $-\mathbf{u}$. The wavefronts of such a wave are shown in Fig. 1.16. The plane wave is weighted by the data $\chi(\mathbf{u})$ to form the integrand,

$$q_{obj}(\mathbf{r}|k,\theta) = \frac{k}{4\pi^2}\{1 - \sin(\theta)\}\, \chi(k,\theta) \times \exp\{jk[-\xi\cos(\theta) + \eta\{1 - \sin(\theta)\}]\} \ . \tag{1.236}$$

This is the 'back-projection' of the data into the space T_{obj}. The back-projections can be combined with respect to angle or frequency in two

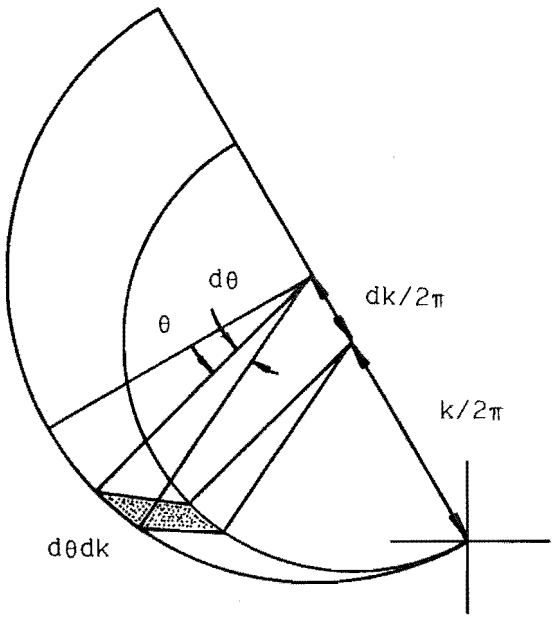


Fig. 1.15 The area element dθdk. The shaded region is the area element of Fourier space corresponding to dθdk.

equivalent ways,

$$q_{obj}(\mathbf{r}|k) = \int d\theta \; q_{obj}(\mathbf{r}|k,\theta) \tag{1.237}$$

and

$$q_{obj}(\mathbf{r}) = \int dk \; q_{obj}(\mathbf{r}|k) \tag{1.238}$$

or

$$q_{obj}(\mathbf{r}|\theta) = \int dk \; q_{obj}(\mathbf{r}|k,\theta) \tag{1.239}$$

and

$$q_{obj}(\mathbf{r}) = \int d\theta \; q_{obj}(\mathbf{r}|\theta) \; . \tag{1.240}$$

The second method is implicit in reconstruction techniques which back-project time delay information as in Hanstead (1981).

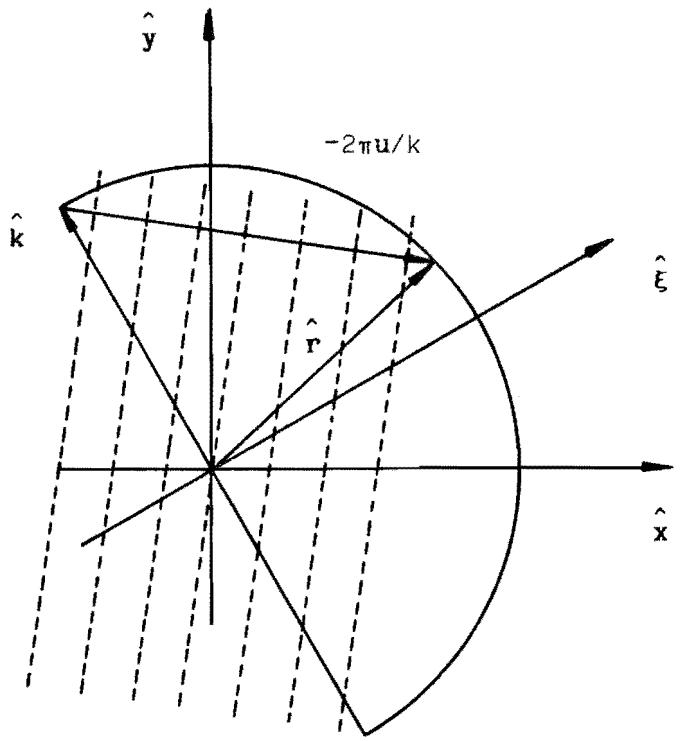


Fig. 1.16 The back-projection interpretation of the inverse Fourier integral. The kernel of the inverse Fourier transform represents a plane wave travelling in the direction $-\mathbf{u}$. The dashed lines indicate the wavefronts of such a wave for a typical observation direction $\hat{\mathbf{r}}$. The reconstruction of the object is a weighted sum of such plane waves projected onto object space.

1.8.5 Back-Projection Methods Based on the One-Dimensional Fourier Transform

The integration in (1.237) is performed for each point of the object. Several of these integrations can be performed concurrently as an FFT if the integral can be expressed as a one-dimensional Fourier transform. This is the case and $q_{obj}(\mathbf{r}|k)$ can be obtained on profiles of constant ξ or η by this method.

Inspection of (1.236) shows that $\cos(\theta)$ and $\sin(\theta)$ are candidates for the role of Fourier transform variable. The transform variable must run from $-\infty$ to ∞ over the domain of integration and should be monotonically related to θ at least over the aperture. Since $\cos(\theta)$ and $\sin(\theta)$ are oscillatory, this last condition cannot be met over the range $[0,2\pi]$ by a single transform variable. As a consequence of this, the methods derived here are suitable only for apertures belonging to one of four classes. For an aperture of the first class, θ lies in the interval $(-\pi,0)$ and for the second, third and fourth classes in $(-\pi/2,-\pi/2)$, $(0,\pi)$, and $(\pi/2,3\pi/2)$ respectively. The equivalent aperture for projection measurements is of the third kind (¶1.7.3). Since the first class includes $\theta = -\pi/2$ which is back scattering it is natural to refer to it as the back scattering class. Similarly the second, third and fourth classes are called the 'right', 'forward' and 'left' scattering classes.

For the left and right scattering classes, the Fourier transform variable is chosen to be

$$u = k \sin(\theta)/2\pi \ . \tag{1.241}$$

The integral (1.224) is then expressed as

$$q_{obj}(\mathbf{r}|k) = -\pm \frac{\exp[jk\eta]}{2\pi} \times \mathcal{F}_u\{w(u)\ \chi(u,k)\ \exp[-jk\xi(1 - \lambda^2u^2)^{1/2})]\}(\eta) \tag{1.242}$$

where $\mathcal{F}_u\{\cdot\}(\eta)$ denotes the evaluation of the one-dimensional Fourier transform of the expression within the braces with respect to u at η. The upper sign in (1.242) is for left scattering and the lower is for right scattering. The minus sign arises because u traverses the aperture in the negative direction, while the integration for the Fourier transform is in the positive direction. The function w(u) is a weighting factor (Fig. 1.17),

$$w(u) = (1 - \sin(\theta))/\cos(\theta)$$

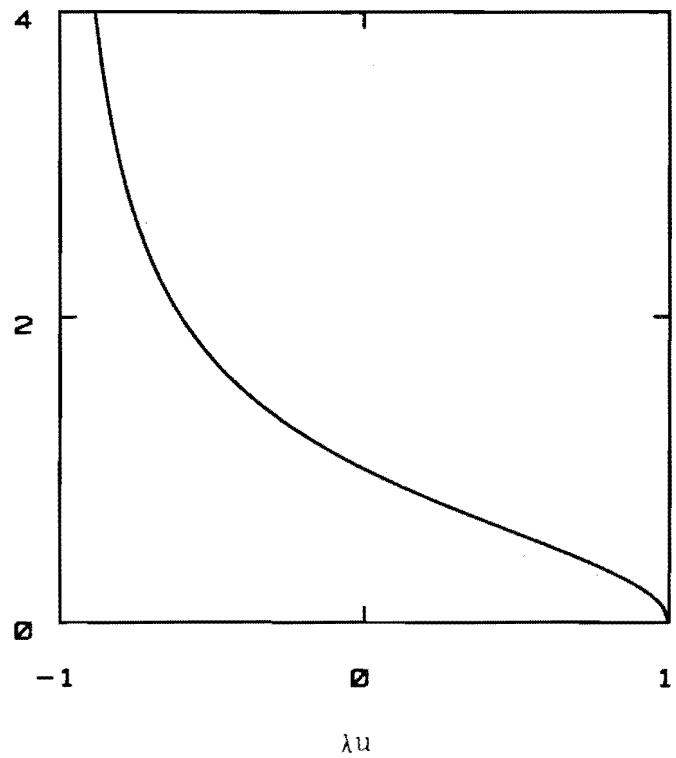


Fig. 1.17 The weighting factor for left and right scattering.

$$= [(1 - \sin(\theta))/(1 + \sin(\theta))]^{1/2}$$

$$= [(1 - \lambda u)/(1 + \lambda u)]^{1/2} \qquad (1.243)$$

which emphasises the back scattering data for the reasons given in ¶1.8.3.

For back and forward scattering the equations corresponding to (1.241), (1.242), and (1.243) are

$$u = k \cos(\theta)/2\pi \ , \qquad (1.244)$$

$$q_{obj}(\mathbf{r}|k) = \frac{\exp[jk\eta]}{2\pi} \times \P F_u\{w(u)\ \chi(u,k)\ \exp[-jk\eta(1 - \lambda^2u^2)^{1/2})]\}(\xi) \qquad (1.245)$$

and

$$w(u) = (1 - \lambda^2u^2)^{-1/2} \pm 1 \ . \qquad (1.246)$$

The upper sign of (1.246) is for back scattering and the lower is for forward scattering (Fig. 1.18).

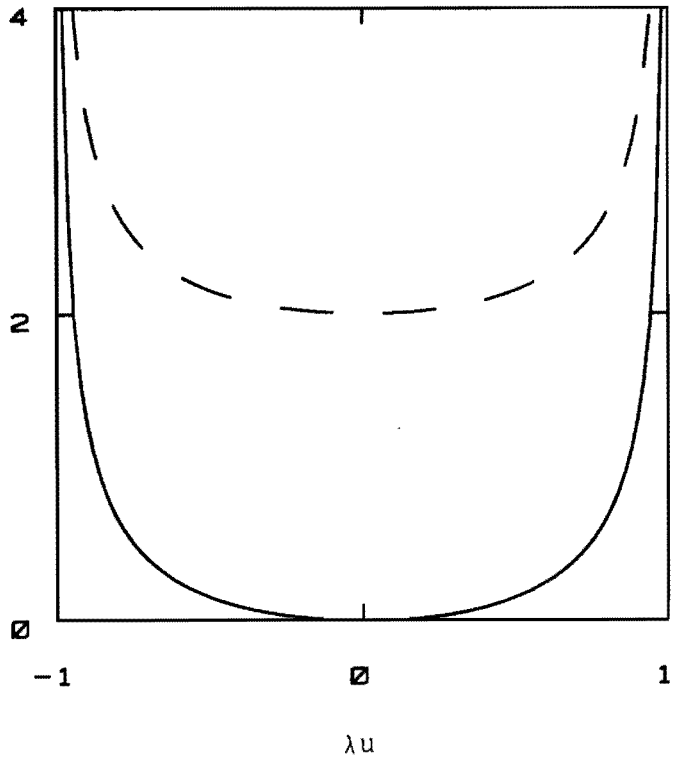


Fig. 1.18 The weighting factor for backward and forward scattering. Dashed line: backward scattering. Solid line: forward scattering.

1.8.6 The Filtered Back-Propagation Algorithm

When the frequency of the incident wave is fixed the integral (1.224) may be expressed with respect to the angle θ and the angle of incidence of the plane wave. In other words, the Fourier space coverage is obtained by varying the direction of incidence and the direction of observation. For the case of projection measurements, Devaney (1982) derives a method for evaluating (1.224) called the 'filtered back-propagation algorithm'. In this method the projection data are filtered and back-propagated from the line $\eta = R$ to the region T_{obj}. The results are

summed for projections obtained at many different angles. The term back-propagation is justified by considering the forward propagation of the phase within the Rytov approximation. A geometric interpretation of the algorithm is presented by Oristaglio (1983). Computer simulations are reported by Devaney (1983) and Pan and Kak (1983).

Pan and Kak (1983) compare filtered back-propagation with interpolation followed by two-dimensional Fourier inversion. They conclude that the latter produces images of comparable quality with considerably less computation. A similar conclusion is reached for the zero wavelength case of CT by Stark et al. (1981). Direct Fourier inversion is shown to be more computationally efficient than filtered convolution back-projection. The latter is the zero wavelength analogue of filtered back-propagation. However, the differences in efficiency for the size of images used in practice is not great.

The conclusions of Stark et al. (1981) derived for X-ray CT probably have some significance for ultrasonic CT. In X-ray CT the projection theorem holds very accurately and so to obtain images which are as accurate as the projection data, the interpolation step must be done very carefully. However, in ultrasonic CT the projection theorem holds only approximately since multiple scattering effects are ignored and so the accuracy of the interpolation step is not so critical. In X-ray CT, the overhead incurred in performing the very accurate interpolation required offsets the efficiency of the Fourier inversion step so that filtered convolution back-projection is preferred. However in ultrasonic diffraction tomography, back-propagation is sufficiently computationally expensive to tip the balance in favour of the direct Fourier inversion method.

1.9 SPECKLE

In practice, an image $\hat{q}_{obj}$ obtained from a CAI system based upon Fourier imaging (¶1.7) usually has a mottled appearance. The amplitude of the image varies throughout the image in a rapid and an apparently random manner. This artefact is called 'speckle'.

In Fourier imaging, measurements of the wave scattered by the object are used to reconstruct an image of the object. The amplitude of a point P in the image depends on the amplitudes of neighbouring points in a resolution cell R of the object. It is assumed that the object is effectively uniform within a resolution cell so that the image amplitude at

P can represent the object over the entire cell. The image is reconstructed by assuming that the wave scattered from P propagates outward as a spherical wave such as the one described by (1.126).

When either or both of the above assumptions are not justified the reconstructed image contains speckle (Abbott and Thurstone 1979). If the object varies rapidly within the resolution cell, 'reflection' speckle results (Bates and Robinson 1981). If the reference medium is inhomogeneous, the wave scattered from P is also no longer spherical and speckle is produced. This is called 'transmission' speckle (Bates and Robinson 1981). If the scattered wave is not measured accurately it is effectively not a spherical wave as far as the reconstruction algorithm is concerned. This is an additional cause of transmission speckle.

1.10 SUMMARY

Several aspects of CAI are reviewed in this chapter. In Part II of this thesis, the results of the author's investigation into several of these aspects is reported.

All of the experimental results reported in this thesis are obtained from ultrasonic scattering measurements. The propagation of ultrasound is discussed in ¶1.3 and methods for calculating the scattering when the form of the object is known are discussed in ¶1.6.

To solve the remote probing problem defined in ¶1.5, an image of an object must be obtained from a data set. The data sets for the results presented in this thesis are obtained with the aid of an ultrasonic scanner. The scanner is described in Chapter 5 using the terminology for CAI introduced in ¶1.5.

An algorithm is required to obtain an image of the object from the data set. When the object scatters weakly, Fourier imaging (¶1.7) may be used. Practical ultrasonic imaging systems based upon Fourier imaging are reviewed in Chapter 2. The images from such systems usually contain the speckle artefact introduced in ¶1.9. An imaging principle, called 'shift-and-add', which is designed to compensate for the speckle artefact, is discussed in Chapter 2. Shift-and-add is described in detail in Chapter 6, where it is also shown that, in conjunction with Fourier imaging, it is an approximate solution to the remote probing problem in a distorting medium.

The mathematical basis for imaging established in ¶1.2 provides

the theoretical framework for the generalisation of shift-and-add which is described in Chapters 7, 8 and 9. The extended imaging principle is a more accurate solution of the remote probing problem. The mathematical basis is also incorporated into the design of the software module for image processing which is described in Chapter 4.

2. ULTRASONIC IMAGING SYSTEMS

2.1 INTRODUCTION

In this chapter, a recent review of ultrasonic imaging systems by Robinson (1982 Ch.3) is summarised and updated. The results reported in the literature are classified according to the imaging algorithm employed and a section is devoted to each. In ¶2.2 the simplest viable scheme, employing the elementary echo-location principle, is described. The conditions under which the echo-location principle is valid are actually quite restrictive. The effectiveness of several echo-location systems in the presence of non-ideal conditions is discussed in ¶2.3.

Systems employing the conventional reconstruction techniques of computed tomography, which are based on the assumption that the wave motion can be represented by rectilinear ray propagation, are reviewed in ¶2.4. The experimental results reported for diffraction tomography are assessed in ¶2.5. Several iterative schemes have been proposed for accounting for ray curvature and these are considered in ¶2.6. A relatively new technique based upon a combination of speckle processing and Fourier imaging is discussed in ¶2.7.

2.2 ECHO-LOCATION METHODS

2.2.1 The Echo-Location Principle

In an echo-location system, the object to be imaged is probed by a narrow beam of ultrasound. Inhomogeneities in the object which are insonified by the beam scatter ultrasound. The reflected component of the scattering wave travels back along the beam and is received by the same transducer used to create the beam. A short pulse of a few cycles duration is transmitted so that the transducer may be used for both transmission and reception. The range of a scatterer is estimated from the time delay between pulse transmission and echo reception. It is assumed that the acoustic velocity is constant throughout the object. The bearing of the scatterer is given by the orientation of the transducer.

2.2.2 Transducers

The resolution of an echo-location system is largely determined by the nature of the transducer even though range resolution is governed by the length of the pulse, which in turn is limited by the system bandwidth. The point is that the system bandwidth is usually restricted by the frequency response of the transducer(s).

The transverse resolution is given by the beam width, which is primarily determined by the transducer aperture. In the far-field (¶1.7.2) of a transducer, the beam diverges. However, in the region between the transducer and the far-field, which is known as the 'near-field' or 'Fresnel zone' (Silver 1965 ¶6.9), the beam width can be arranged to vary from being of the same order as the diameter of the transducer, to being of the order of a wavelength. The reduction of the beam width below the diameter of the transducer is called 'focusing', and may be achieved by special shaping of the surface of the transducer, or by using arrays (Macovski 1976). The range over which the transducer is considered to be focused is called the 'depth of focus' (Wells 1977 p.152). Strong focusing, which produces a very narrow beam width, results in a small depth of focus. Weak focusing (Kossoff et al. 1968) is usually employed to allow the imaging system to function over a wider range without refocusing. The near-field transient response of transducers and arrays of transducers is investigated by Robinson et al. (1974), Weyns (1980), Hefner et al.(1980), Arditi et al. (1981), and Duerinckx (1981).

The transverse resolution is also affected by the system bandwidth. Although the width of the main lobe is not a strong function of bandwidth (Miller et al. 1974), the side lobes can affect the transverse resolution. Consider the case when several point scatterers are close together, but are separated by more than the width of the main lobe. The side lobes may make it difficult to determine the location of the scatterers (Eaton et al. 1980). Narrow-band operation tends to result in strong side lobes. As the bandwidth is increased the effective side lobe levels decrease (Hildebrand 1980).

Side lobe levels can be controlled by weighting various sections of the aperture differently. This is referred to as 'apodisation' (Prabhu et al. 1977; Harris 1978). Acoustic lenses are often naturally apodised (Dunn and Fry 1980). The weighting arises because the lenses are usually concave and made of a material more attenuating than the medium in which they are used. When a computer is used for image reconstruction, an alternative to

apodisation known as 'edge extension', which involves less resolution loss, may be used (McDonnell and Bates 1975).

2.2.3 Echo Signal Processing

The signal-to-noise ratio of the echo signal may be increased by boosting the transmitted power. The amount by which the amplitude may be increased is limited, however, by the breakdown characteristics of the transducer, the medium, and the object. An alternative method of increasing the effective power is to use longer duration signals than pulses. Phase encoding is required to maintain the system bandwidth when signals of long duration are employed (Sato and Wadaka 1979, Chapelon et al. 1979). The received signal is correlated with the transmitted signal to recover the time delay information. The procedure is known as 'matched filtering' or 'pulse compression' (Rihaczek 1969 ¶2.5, Takeuchi 1979).

Several techniques have been proposed for extracting time delay information from the filtered signal. The time of first arrival is commonly used. Gamell (1981) suggests a sophisticated method in which the analytic signal (Gabor 1946; Cook and Bernfeld 1967 ¶4.2) is computed from the filtered signal.

2.2.4 Displays

The echo signal received by the transducer encodes the reflectivity of the object along the beam as a function of time. To display the reflectivity, several other signals must be generated by the imaging system. A 'range' signal indicates the range as a function of time and is generated by triggering a ramp waveform generator every time a pulse is transmitted. Other signals may be generated to indicate the location of the beam. Several techniques are available for displaying the reflectivity image on an oscilloscope (Reintjes and Coate 1952 ¶4.4).

An 'A-mode' display is obtained by deflecting the spot vertically with the echo amplitude and horizontally with the range signal (Robinson et al. 1970; Wells 1977 ¶6.7). In 'B-mode', the spot is intensity modulated by the echo amplitude and deflected by the range signal (Wells 1977 ¶6.8). A 'C-mode' display is produced by setting the spot intensity according to the echo amplitude at a particular range and deflecting the spot to a position representing the location of the beam (Wells 1977 ¶6.10a). When the echo signals from successive pulses are displayed one beside another in B-mode, the result is called an 'M-mode' display (Macovski 1983 p.177).

2.2.5 Scanning

To obtain reflectivity information over a two-dimensional slice of the object, the beam must be scanned. The principles of the two basic methods, linear and sector scanning, can be conveniently explained with reference to Figs. 2.1(a) and (b). In linear scanning the transducer is moved perpendicularly to the direction of the beam. Sector scanning is achieved by rotating the transducer about a fixed location so that the beam is swept over a range of angles. Compound scanning, illustrated in Fig. 2.1(c) is a combination of linear and sector scanning. When an image is reconstructed using the echo-location principle, the finite beam width is ignored and the beam is assumed to be a 'scan line' through the image.

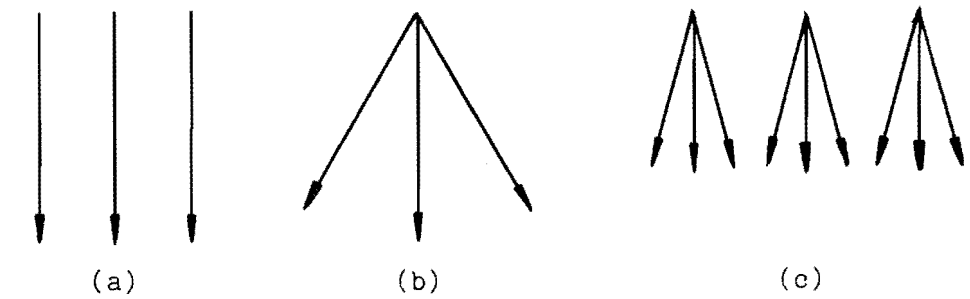


Fig. 2.1 Scanning techniques. (a) linear. (b) sector. (c) compound.

The process of combining the echo signals obtained from different beam locations is called 'scan conversion'. Both digital and analog devices are available for scan conversion. Digital devices have a greater dynamic range and afford more versatile processing (Ophir and Maklad 1979). A digital scan convertor is essentially an array of digital memory. A typical memory size is 512 × 512 pixels quantised to 4 or 5 bits.

The major difficulty associated with digital scan conversion arises from the interaction of the scan lines and the pixel positions. Some pixels are not crossed by any scan lines and so do not receive any echo information. This leads to a moiré effect. Algorithms to fill these 'holes' are discussed by Ophir and Maklad (1979). When two scan lines intersect a 'writing algorithm' is invoked to combine the associated echo

amplitudes. The usual method is to use the maximum amplitude (Robinson and Knight 1981), but the minimum and average amplitudes offer better resolution when the object has a uniform refractive index. The names of these algorithms are 'maximum', 'minimum', and 'average' amplitude writing, respectively. This is discussed further in ¶2.3.2. A recent discussion of scan conversion is that by Robinson (1984).

2.2.6 Image Processing

Image processing may be employed to aid the interpretation of the reflectivity image. Procedures such as image expansion, histogram equalisation, edge enhancement, smoothing, windowing, thresholding and pixel filling to eliminate moiré patterns (Bates and McDonnell 1985; Ophir and Maklad 1979) are useful. Such procedures are useful because they can make the information content of the reflectivity image easier to display and because they can remove artefacts (¶1.5.3). For instance, the diagnosis of breast disease from reflectivity images is facilitated by image processing (Jellins et al. 1975; Kossoff et al. 1976; Kobayashi 1979).

Deconvolution and Wiener filtering can increase the resolution by taking into account the finite size of the focal region (Hundt and Trautenberg 1980). Such techniques assume a model for the reflectivity image r of an object f of the form

$$r = f \odot h + c \qquad (2.1)$$

where h is called the 'point spread function (PSF)' and c is called the 'contamination (function)'. However, an accurate knowledge of the PSF is required and the procedure is sensitive to noise. Vakine and Lorenz (1984), Robinson and Wing (1984) and Schomberg et al. (1983) investigate methods for improving the transverse resolution of B-mode images.

2.2.7 Arrays

An 'array' is a configuration of discrete transducers. Arrays permit mechanical scanning movements to be replaced by electronic switching. Medical imaging systems based upon arrays are able to operate at rates of up to fifteen images per second. This limit is imposed not by the imaging device but by the time required for ultrasound to propagate within the object. Some of the medical imaging systems employing array techniques which are in routine clinical use are described by Ligtwoet et al. (1978), Morgan et al. (1978), Selbie et al. (1980) and Assenza and

Pappalardo (1980). Development systems employing array techniques are described by Karrer et al. (1982) and Peterson and Kino (1984).

Electronic Scanning

An array consisting of narrow-beam transducers may be used to eliminate the mechanical scanning required when a single transducer is used for linear or sector scanning (Bom et al. 1973). A more powerful approach, however, is to use the array to act as a single transducer with a controllable aperture.

Fig. 2.2 illustrates the essential features of a receiving array with a controllable aperture consisting of a number, say N, of array elements. The i^{th} array element consists of a wide beam transducer T_i and a device D_i, called a 'delay element', for delaying the signal from T_i. The beam width of T_i is sufficient for the transducer to respond to an echo $p(t)$ from an arbitrary point P in the object. The propagation delay from P to T_i is denoted by $\tau_i^{(P)}$ and the time delay introduced by D_i is $\tau_i^{(D)}$. The output $s(t)$ of the array is the sum of the outputs $s_i(t)$ from each delay element,

$$s(t) = \sum_{i=1}^{N} s_i(t) \ . \tag{2.2}$$

The total time delay $\tau_i^{(S)}$ from P to the output of D_i is given by

$$\tau_i^{(S)} = \tau_i^{(P)} + \tau_i^{(D)} \ , \tag{2.3}$$

hence $s_i(t)$ can be written as

$$s_i(t) = p(t - \tau_i^{(S)}) \ . \tag{2.4}$$

The output of the array is maximised if all of the delays $\tau_i^{(S)}$ are equal, i.e.

$$\tau_i^{(S)} = \tau \ , \quad 1 \leq i \leq N \ . \tag{2.5}$$

This is achieved by controlling the delay elements so that

$$\tau_i^{(D)} = \tau - \tau_i^{(P)} \ . \tag{2.6}$$

The propagation delay $\tau_i^{(P)}$ is calculated from the positions of P and T_i and from the constant acoustic velocity assumed for the medium. The constant time delay τ is chosen to be greater than or equal to the largest propagation delay. With each of the time delays $\tau_i^{(D)}$ set according to

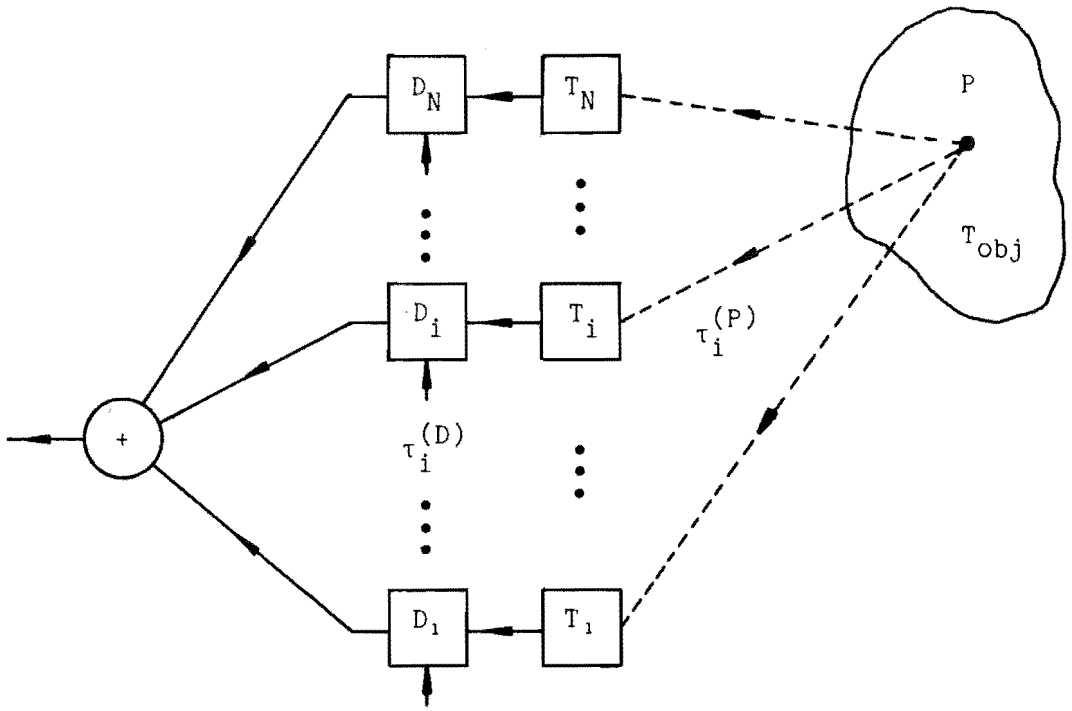


Fig. 2.2 An electronically scanned receiving array. The array consists of N array elements. Each element consists of a delay element D_i, which introduces a time delay $\tau_i^{(D)}$ into the received signal, and a transducer T_i, which receives the echo from a point P after a propagation delay of $\tau_i^{(P)}$. The outputs from each delay element are summed to produce the array output.

(2.6), the array output is given by

$$s(t) = N\, p(t - \tau) \tag{2.7}$$

and the array is focused at P. By altering the delays electronically, the beam may be rapidly scanned over the entire object (Thurstone and von Ramm 1974).

Beam Pattern

The aperture of an array can be considered to be a sampled version of the continuous aperture of a single large transducer. Signals which are not sampled closely enough cannot be reconstructed because of an effect known as 'aliasing' (Bracewell 1978 ¶10). An analogous difficulty can affect the beam pattern of an array. Consider the situation depicted in Fig. 2.3. The array is steered in the direction P and is operating with

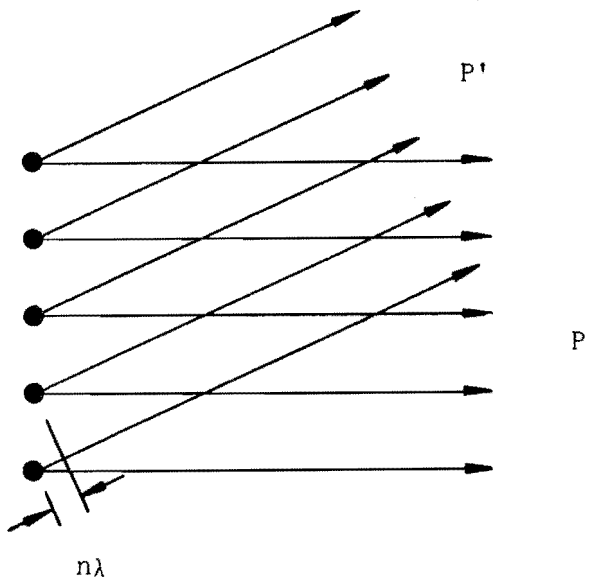


Fig. 2.3 Grating lobes. The array is operating with radiation of wavelength λ and is steered toward P. The direction P' is in a grating lobe of the array because the indicated path difference is a multiple of λ.

monochromatic radiation of wavelength λ. However, for paths in the direction P', originating from adjacent array elements, the difference in path length is an integral multiple of the wavelength. Radiation from P' is therefore in phase for all elements of the array and makes a significant contribution to the output of the array. Radiation in any such direction is known as a 'grating' lobe. Steinberg (1976 ¶5.2) shows that the

element spacings in an array must be less than one half of a wavelength to avoid grating lobes. When wide bandwidth signals are employed larger spacings can be used because the position of the grating lobes is frequency dependent (von Ramm and Smith 1978; Hildebrand 1980). Grating lobes may also be reduced by spacing the array elements randomly (Steinberg 1976 ¶8.1). Steinberg and Luthera (1982) describe methods of reducing the side lobes of a ring-shaped array.

Integrated Arrays

Considerable research is being conducted into 'integrated' arrays. These are arrays fabricated on a single substrate using microelectronic technology. Piezoelectric materials such as ceramics (Plummer et al. 1978) and polyvinylidene fluoride (Swartz and Plummer 1980) can be used in an integrated array. The piezoelectric material forms the gate electrode of a field effect transistor.

The substrate connecting the elements can cause difficulties. Acoustic coupling between elements may occur and the sensitivity may be less than for an array consisting of discrete elements. This happens because the acoustic impedances of the materials used may not be optimal. The close spacing employed in integrated arrays can cause electrical coupling (Bruneel et al. 1979), and the small size of each element can lead to coupling to transverse vibrational modes (Smith et al. 1979).

Delay Elements

Delay elements can be constructed from analog or digital components. Switchable wide-band lumped-constant analog circuitry is used by Thurstone and von Ramm (1974). Maginess et al. (1976) describe a sampled analog method employing a charge-coupled device (CCD). A digital technique using a shift register is described by Duck et al. (1977). Tapped surface acoustic wave (SAW) delay lines allow the delays required for electronic scanning to be selected easily. The delays are determined by the waveform propagating in the delay line. Cauvard and Harteman (1982) describe an ultrasonic imaging system using SAW components which can generate images containing 1500 pixels at a rate of 800 images/second. The design of SAW devices is discussed by Engan et al. (1984).

Configurations

The configuration of the array elements is chosen to yield a beam with a narrow main lobe and low side lobes, subject to cost and construction constraints.

'Sparse' arrays, consisting of very few elements, are attractive from the point of view of cost. When separate transmit and receive arrays are used, the combined beam pattern is the product of the transmit and receive patterns. It follows that the beam pattern of two sparse arrays, which by themselves may have unacceptably high side lobe levels, may be combined as a transmit-receive pair with a satisfactory beam pattern (Kossoff 1973; Fraser et al. 1975; Nitadori 1975; Nitadori et al. 1980).

Arrays with axial geometry have simpler construction requirements than those of arbitrary geometry. Burckhardt et al. (1974a), Vilkomerson and Hurley (1975), Macovski and Norton (1975), and Alais and Richard (1980) describe imaging systems with transducer elements in the form of concentric annuli. Patterson and Foster (1983) and Foster et al. (1983) describe an unusual annular array - conical transducer configuration.

2.3 ECHO-LOCATION IN PRACTICE

2.3.1 Overview

The simple acoustic model of the medium and object, implicit in the echo-location principle (¶2.2.1), neglects attenuation and the possibility of acoustic velocity variations within the object. However, the imaging systems described in this section may often be configured or designed in such a manner as to lessen the practical limitations inherent in these simplifications.

2.3.2 Refractive Index Variations

Variations in the refractive index imply that the propagation occurs along curved ray paths. The positions inferred from the echo-location principle are incorrect if this is the case. A more serious consequence is that the scan conversion procedure of ¶2.2.5 cannot combine the echoes from a point of the object if they are obtained by insonifying the point from different directions. This error is called 'mis-registration'.

Consider the image of a point scatterer P which is subject to mis-registration. As the beam is scanned through the object, P is insonified several times. However, because of mis-registration the corresponding scan lines do not all intersect at the same place. So the same scatterer tends to be imaged at several different points, albeit probably closely spaced. Minimum amplitude writing (¶2.2.5) results in most, if not all, of the points vanishing, while maximum amplitude writing allows each point to contribute to the image. The latter writing method is therefore more tolerant of mis-registration. If the image obtained by maximum amplitude writing is smoothed (¶2.1.6), the collection of points is transformed into a single region. The result is a lower resolution image than would have been obtained in the absence of mis-registration.

Smith et al. (1984) reduce mis-registration by modifying the echo-location principle. They assume that the beam travels through a region whose refractive index is known and constant, and that the extent of the region is known. With this knowledge, the path of the beam as it travels through the region may be calculated. Echoes can therefore be better assigned to the point from which they originated so that mis-registration is reduced.

Robinson et al. (1983) take advantage of mis-registration to estimate the refractive index. The object is assumed to have a uniform acoustic velocity and ray calculations are used to deduce the velocity from the observed mis-registration.

2.3.3 Attenuation

Attenuation is not taken into account in the echo-location principle described in ¶2.2.1. In practice, its effects are accommodated in an ad hoc manner (Maginness 1979). The usual approach is based upon the assumption that the attenuation coefficient is constant throughout the object. This implies that the attenuation is proportional to range and can be allowed for by boosting the amplitude of the received signal by an amount proportional to time. This is known as 'time gain compensation (TGC)'. A more sophisticated method is discussed by De Clercq and Maginess (1975). They allow the attenuation coefficient to vary with range and propose a method for estimating it from the echo signal itself.

Attenuation of a wave as it propagates through a medium is due to absorption (Longley and O'Brien 1982), scattering (Pohlhammer and O'Brien 1981), and diffraction (¶2.2.2). Absorption is the conversion of the

energy of the pulse into heat. The loss due to diffraction is associated with the spreading of the pulse as it propagates. Methods for estimating attenuation from echo-location signals are discussed by Leeman et al. (1984), Kuc (1984), and Robinson et al. (1984).

2.3.4 Absorption

In theory, higher resolution is obtained by operating at higher frequencies. However, for many media, including most soft tissues, absorption is proportional to frequency. The proportionality constant is called the 'frequency dependent attenuation coefficient'. For small inhomogeneities this effect may be partially offset by increased scattering at higher frequencies. Morse and Ingard (1968 ¶8.2) show that the amplitude of the wave scattered from such an inhomogeneity is proportional to the square of the frequency. In general though, as the frequency is increased, the echo strength is reduced and eventually, beyond some maximum range, becomes lost in noise. Quite apart from the irreducible thermal noise, extra noise arises in the imaging system itself and from various sources in the medium. Typically, upper frequencies are 15 MHz for opthalmology and 5 MHz for abdominal investigations.

2.3.5 Scattering

Consider the object shown in Fig. 2.4 which consists of two scatterers, each of which has linear dimensions of several wavelengths. Furthermore, suppose the object is probed by a beam of ultrasound which passes through both scatterers. A pulse incident upon the object is partly scattered and partly transmitted by the scatterer it encounters first. The transmitted pulse then encounters the second scatterer and is once again partly scattered and partly transmitted. The scattered pulses give rise to echoes from each scatterer. However, the echoes are not proportional to the reflectivities for each scatterer because the pulse that strikes the second is weaker than the pulse that strikes the first. In the reflectivity image the second scatterer appears to be shadowed by the first. This 'shadowing' artefact may be ameliorated by using the compound scanning technique described in ¶2.2.5. During the course of compound scanning, beam positions are used for which the pulse strikes the second scatterer without travelling through the first. The resultant echoes are proportional to the reflectivity of the second scatterer and help to offset the shadowing artefact.

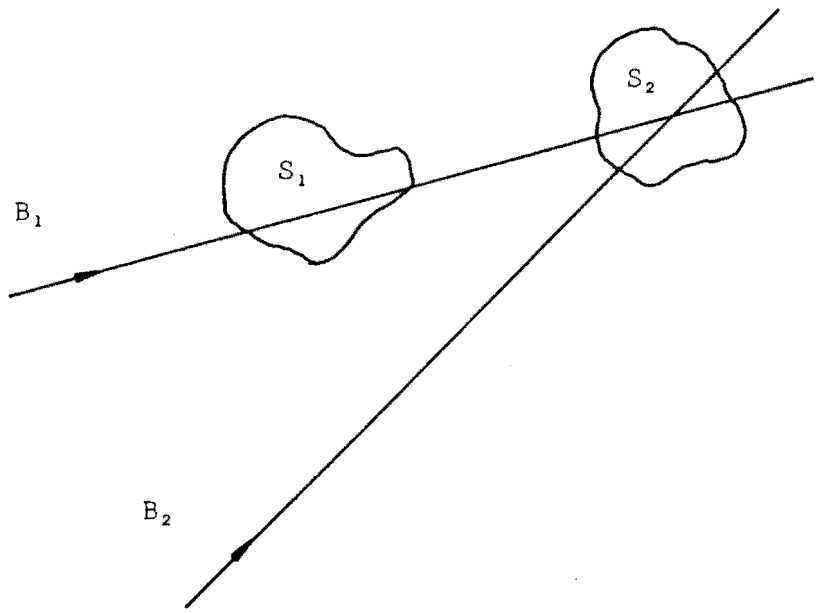


Fig. 2.4 The shadowing artefact. For the beam position B_1, the scatterer S_2 is in the shadow of the scatterer S_1. In compound scanning, beam positions such as B_2 are also used and in this case S_2 is probed directly.

2.3.6 Specular reflections

Specular reflections occur at smooth interfaces between regions of differing acoustic impedance. A (back scattered) echo is received from such an interface only at normal incidence. A linear or sector scan (¶2.2.5) thus tends to reveal only a small part of the interface. However, during a compound scan echoes are obtained from an appreciably larger part of the interface.

It is worth noting that the image of a point scatterer may be obliterated if it is positioned close to a smooth interface. A strong specular reflection from the interface contributes to the reflectivity image at the location of the interface and at some of the pixels nearby. The former contribution is due to the main lobe of the beam, the latter to the side lobes. It is possible that the image of the point scatterer may be

swamped by the contribution from the side lobes.

Minimum amplitude writing (¶2.2.5) may enable the point scatterer to be distinguished (Linzer and Parks 1979). The point scatterer gives an echo for any beam location which passes through the point and therefore contributes to the reflectivity image. However, the specular reflection from the interface occurs at only a few positions and so does not contribute significantly.

2.3.7 Speckle

The speckle (¶1.9) present in a reflectivity image obtained by linear or sector scanning may be reduced by using a compound scan and maximum amplitude writing (Burckhardt 1978). However, a more significant reduction is obtained by using an incoherent average to combine the echoes from intersecting scan lines (Wells and Halliwell 1981; Robinson and Knight 1981). Melton and Magnin (1984) present a model for predicting the degree of speckle reduction which accounts for the influence of system bandwidth.

2.4 COMPUTED TOMOGRAPHY

2.4.1 Attenuation

Projections of the attenuation coefficients are obtained by transmitting a narrow beam of ultrasound through the object to be imaged and comparing the incident and transmitted amplitudes. Greenleaf et al. (1974) uses CT (¶1.7.1) to reconstruct images of attenuation. The results have inferior resolution to X-ray CT. The reason for this is that (as discussed in ¶2.3.3) attenuation is due to three factors which cannot be easily separated. X-ray CT is successful because two of them, scattering and diffraction, can usually be neglected.

Refraction can reduce the accuracy of estimates of the attenuation coefficient obtained by ultrasonic CT. During reconstruction, an effect similar to mis-registration (¶2.3.2) can occur. Other difficulties arise during the measurement of attenuation. The beam may be bent by refraction effects to such an extent, that the main lobe misses the receiving transducer altogether. Carson et al. (1977) suggest the use of a large receiving transducer to overcome this difficulty. However, even if the beam does strike the transducer, an interference effect called 'phase cancellation' may occur. It arises when the beam strikes the receiving transducer at other than normal incidence. In this case, the phase of the

beam varies across the receiving aperture and cancellation occurs if the transducer is phase sensitive. Klepper et al. (1981) suggest the use of phase insensitive transducers to overcome this, but such transducers tend to have low sensitivity. Schmitt et al. (1984) use a large receiving array and incoherent signal processing to avoid phase cancellation.

Absorption is frequency dependent (¶2.3.4) and this can make wide-band attenuation measurements difficult. Dines and Kak (1979) use a frequency shift method to obtain the 'integrated' attenuation coefficient. This method is based upon the change in spectral content of a wide bandwidth pulse as it propagates through the object. An advantage of the method is that it is insensitive to variations in the absolute signal level caused by reflection and refraction.

2.4.2 Refractive Index

CT reconstructions of the refractive index are described by Greenleaf et al. (1975), Bates and Dunlop (1977) and Glover and Sharp (1977). Projections of the refractive index are obtained by time delay measurements (¶1.7.1).

Multipath effects can occur. Dunlop (1978 ¶3.3) suggests that most of the energy is transmitted along the direct path. Consequently, time delay measurements based upon the maximum amplitude of the received pulse are likely to be more accurate than those based on first arrival times (¶2.2.3). Provided multipath effects are identified, projections of the refractive index may be obtained more accurately than those of attenuation. McKinnon et al. (1984) ignore projection data if multipath effects can not be identified sufficiently confidently to determine the signal corresponding to the direct path. Crawford and Kak (1981) investigate the use of median or homomorphic filtering to compensate for multipath artefacts. Although both techniques are found to be equally effective, median filtering is favoured because it requires less computation.

2.5 DIFFRACTION TOMOGRAPHY

2.5.1 Multiple Scattering

The theory of diffraction tomography is presented in ¶1.7.1. This theory is approximate in two respects. Firstly, the wave equation (1.112), upon which diffraction tomography is based, is only an approximate model for the interaction of ultrasound and the medium. Secondly, only first

order scattering is considered since either the (first) Born or the (first) Rytov approximations are invoked.

The adequacy of the Born and Rytov approximations in the context of diffraction tomography is discussed by Azimi and Kak (1983). Their discussion is couched in terms of multiple scattering. The neglect of multiple scattering is equivalent to using a first order scattering approximation, such as the Born or Rytov approximations. Firstly, they consider the forward scattering problem (¶1.6.1) to determine when multiple scattering is likely to be significant. Secondly, they investigate the artefacts which result when diffraction tomography is used to reconstruct objects for which multiple scattering is significant.

In their investigation of the forward problem they consider objects consisting of several components. Each component has a simple shape for which an exact scattering solution is known. The total scattering from the object consists of the scattering of the incident wave by each component and of multiple scattering between components. This is known as Twersky's formulation of multiple scattering and is discussed by Ishimaru (1978 Ch. 14). Azimi and Kak (1983) conclude that even when the components are weakly scattering, it may not be possible to neglect the effect of multiple scattering if the components are in line with the direction of insonification. When the components are strongly scattering, multiple scattering is always important unless the distance between components is large, or if the components are large and not blocking each other. They illustrate the effect of multiple scattering on the reconstruction of a two component object. Even when the acoustic properties of the object differ from those of the reference medium (¶1.3.5) by as little as 5%, severe distortions can result. The distortions seem to be particularly pronounced along lines which intersect both components. These conclusions are compatible with the intuitive reasoning expounded in ¶2.3.5.

2.5.2 Attenuation

It is explained in ¶1.7.3 that the two-dimensional spatial Fourier transform of the object may be sampled over a two-dimensional region by either varying the frequency or the direction of the insonification. In terms of the theory of Fourier imaging the two methods are equivalent. In practice, they are not because the absorption is frequency dependent (¶2.3.4). In effect, the object imaged by the variable frequency method is an average over frequency. Miyashita et al. (1980) and Yamamoto and Aoki (1984) use the variable frequency method to

obtain acoustical images. The advantage of diffraction tomography using the variable direction method is that the absorption may be obtained at a particular frequency. Greenleaf and Chu (1984) obtain attenuation images at several frequencies using the Rytov approximation and calculate an image of the frequency dependent attenuation coefficient (¶2.3.4).

2.5.3 A Comparison of Computed Tomography and Diffraction Tomography

Given that the propagation model of (1.112) is approximate, one might ask whether the additional complications of diffraction tomographic reconstructions with respect to those of CT are warranted. Robinson and Greenleaf (1985) present results which suggest that they are, at least for some objects. Greenleaf et al. (1982) also illustrate the effects of diffraction on CT. Adams and Anderson (1982) compare diffraction tomography and CT.

Robinson and Greenleaf (1985) present acoustic velocity and attenuation images of a simple test object consisting of a thin wire and two 18 mm diameter rubber tubes. One tube is filled with alcohol and the other with saline solution. Since most of the object is composed of inviscid fluids, (1.112) should be a good approximation. The images reconstructed by diffraction tomography are considerably more faithful than those obtained by CT.

Robinson and Greenleaf (1985) employ the Rytov approximation for their reconstructions by diffraction tomography. Experimentally, the phase is always measured modulo 2π and so must be subject to a 'phase unwrapping' procedure before the complex phase difference (1.194), required by the Rytov method, can be calculated. However, the phase of the wave scattered from the simple test object does not vary by more than 2π radians, hence, phase unwrapping, which is a difficult procedure, is not required. Nahamoo and Kak (1982) avoid the Rytov approximation because of the difficulty of phase unwrapping. Phase unwrapping methods are discussed by Kaveh et al. (1984) and Greenleaf et al. (1984).

2.6 CURVED RAY TECHNIQUES

In general, ultrasound propagates along ray paths which are curved rather than straight. Conventional CT algorithms are based on the assumption that the rays are straight. Diffraction tomography within the Born or Rytov approximations assumes that the scattering is weak. Since strong forward scattering is responsible for refraction and hence ray

curvature, diffraction tomography does not account for ray curvature either.

All of the methods proposed for accounting for ray curvature are iterative techniques of unproven convergence. An initial estimate of the refractive index is obtained by assuming that the ray paths are straight. The actual ray paths are then estimated by applying ray tracing procedures to the estimated refractive index. A new estimate of the refractive index is obtained with a reconstruction technique using the estimated ray paths.

The 'algebraic reconstruction technique (ART)' of Gordon et al. (1970) is more suitable for these iterative schemes than modified back-projection (Greenleaf et al. 1979). This is because the estimates of the curved ray paths are more easily incorporated. An efficient implementation of ART called 'simultaneous ART (SART)' is described by Andersen and Kak (1984). Refraction corrections for objects in which the variations in refractive index are no more than 10% do not seem to be particularly helpful (McKinnon and Bates 1980; Farrell 1981; Johnson et al. 1975; Schomberg 1978, 1982).

McKinnon et al. (1984) describe a scanner designed explicitly for ultrasonic transmission imaging of the female breast. An iterative reconstruction procedure taking into account ray curvature in three dimensions is used in conjunction with the scanner. McKinnon et al. (1984) suggest that the low resolution of ultrasonic CT and diffraction tomography images is due to imperfect measurements. To this end, their scanner gathers projection data rapidly with two continuously rotating linear arrays and employs small transducers to avoid the phase cancellation described in ¶2.4.1.

2.7 SPECKLE PROCESSING

The effects of speckle, which are discussed in ¶1.9, may be allowed for by an image reconstruction procedure in two ways. The simplest approach is to treat speckle as noise and to attempt to average out its effects. This is the viewpoint taken in ¶2.3.7. The alternative is to treat the problem of removing speckle as a blind-deconvolution problem (Stockham et al. 1975). This approach is taken by Robinson (1982).

He applies speckle processing techniques developed for optical astronomy to ultrasonic imaging. He concludes that narrow-band ultrasonic images formed through mammalian soft tissues, specifically liver, are

severely degraded by transmission speckle (¶1.9). Speckle processing is shown to dramatically improve the faithfulness of such 'speckle' images. Speckle processing requires an ensemble of images of the object. The distortions should be statistically independent and the contaminations should not dominate the isoplanatic component (¶1.5.4). Robinson (1982) shows that such an ensemble can be obtained by spectral decomposition.

The objects encountered in diagnostic medicine consist of various types of tissues. Often just a small part of the object, called the 'target', is germane to the diagnosis. For instance, in oncology, the target is a tumour. The information required from the image is firstly whether the tumour is present and secondly, if it is present, whether it is malignant or benign. Hence, an image of the target, rather than of the object as a whole, is required.

The objects investigated by Robinson (1982) are restricted in several respects. Firstly, all of the targets are composed of isolated point scatterers. Extended targets are not considered. Secondly, all of the targets contain a single bright point. The 'coherent shift-and-add' principle proposed by Bates and Robinson (1981) (and discussed in Chapter 6) is only applicable to such objects. Thirdly, the tissue is constrained to be between the target and the measurement aperture. The case in which the insonifying beam travels through tissue before striking the target is not investigated. Also no examples in which the tissue actually touches the target are presented by Robinson (1982).

In Part II of this thesis, the speckle processing proposed by Robinson (1982) is generalised to permit a wider class of targets to be imaged. Additionally, in ¶6.8, the coherent shift-and-add principle is interpreted as an approximate method for solving the remote probing problem (¶1.5).

3. SOFTWARE DEVELOPMENT

3.1 INTRODUCTION

As explained in ¶1.5, CAI is a solution to the remote probing problem. To implement a particular imaging algorithm in a CAI system, the digital computer involved must execute a representation of the algorithm in machine code. In a research enviroment, new algorithms are developed and expressed in machine code so that they may be tested. The effort required to perform this programming step can be considered to be channelled into two activities. The first is devising a way to program the new algorithm and attempting to code it. Since the new algorithm is usually a modification of an existing one, this step consists of software modification. The second step consists of correcting the mistakes made in the first step. This chapter discusses various aspects of these steps in some detail.

In ¶3.2 the relationship between software specification and software modification are examined. Methods for testing software are discussed in ¶3.3. Finally, a test tool which facilitates program testing is proposed in ¶3.4.

A large amount of software was written to obtain the results presented in this thesis. The methods described in this chapter were used to facilitate the development of this software.

3.2 SOFTWARE SPECIFICATIONS

In this chapter the way in which software is created in response to a specific problem is considered. It is assumed that a specification for the software already exists and attention is confined to the successful implementation of the specification in software. The nature of the specification may differ considerably from one project to the next. In one case the requirements might be quite well defined; in another they may be intentionally imprecise, possibly in anticipation of later changes of an unknown nature. Typically, the specification states what actions the software is to perform on some data. Performance constraints such as code size and execution speed are often included. Details of external interfaces to be used may also be present. Such interfaces may be in the form of existing hardware or programs which are to be incorporated into the implementation. These requirements are the most obvious components of a

software specification. The purpose of this chapter is to discuss two other often neglected and less tangible aspects of the implementation, software testing and software modification. It is suggested that tools more advanced than those currently available are required to facilitate software testing.

3.3 A SIMPLE DEVELOPMENT STRATEGY

Traditional software development is associated with ad hoc coding methods and considerable and unpredictable time spent removing mistakes. This is unsatisfactory from a project scheduling point of view and is expensive. For this reason some effort should be expended towards developing an optimal development strategy.

The cheapest strategy for which the end product satisfies the specification should be adopted, e.g. the strategy which produces the quickest solution with the available labour. Unfortunately, a quick implementation is often assumed to follow from writing as little software as possible. Two problems can be identified with this 'minimal' software strategy. One problem is that the software cannot be tested thoroughly. This occurs when intermediate results, generated and used by the software, are not output. Additional software is required to output these results. In a minimal software strategy, however, this is omitted and only the input to the software and the final output are available. If the software has a mistake in it, the output must be incorrect for some inputs. All of the software has to be checked because no intermediate results are available to indicate the section of the software containing the bug. Hence, an elusive bug can delay completion unpredictably and for substantial periods of time. The delay may be of the same order as the time required to write the software in the first place because all of the software has to be inspected.

Another problem is that simple generalisations of the software, which would allow it to handle more cases than are strictly required by the specification, are not made. Later on, if the specification is expanded to include these cases, large parts of the software may be to be reworked. Thus, in the long term, the minimal software strategy may not be the cheapest.

3.4 SOFTWARE MODIFICATIONS

3.4.1 Software Maintenance

While the software specification is ostensibly fixed, in practice it may be altered from time to time. This can happen after an implementation is completed or even during the development of the software. The latter case occurs, for instance, when hardware is being developed in parallel with software and it is realised that there is a simpler way to do things. The requirements of the people using the software usually change with time and so new versions of the software are required to meet updated specifications. Modifying the software to meet the new specification is called 'software maintenance'. Software maintenance is important if software is to remain current for much longer than a year.

It may seem implausible to suggest catering for changes in the specification when the nature of the changes is unknown. However, certain coding practices can ease the burden of software modification considerably. The required modification usually has a simple description in English since simple ideas are more likely to be promulgated. Consequently, a feature of efficient software design and good coding practice is that the transcription of the modification into code should also be simple. This can be achieved by a combination of two methods. One is concerned with the naming convention used when the software is coded. The other pertains to the structure of the specification.

3.4.2 Structured Specifications

Software modification is eased when the structure of the software and of the specification resemble each other. The structure envisaged for the specification is illustrated in Fig. 3.1. It results from dividing the specification into independent 'sections' and then logically ordering the 'requirements' in each section. The parts of the specification should be as independent as possible. One way in which this may be achieved is to have each section deal with a different part of the data. A second method is to put all of the specifications concerned with a particular action into one section. The requirements are ordered so that each relies for its definition only on previously specified requirements. When this is done the requirements at the beginning of each part are more general than those following and so are less likely to be altered when the specification is updated. These 'basic' requirements should be implemented as a library of routines. The rest of the requirements can be implemented by one or more

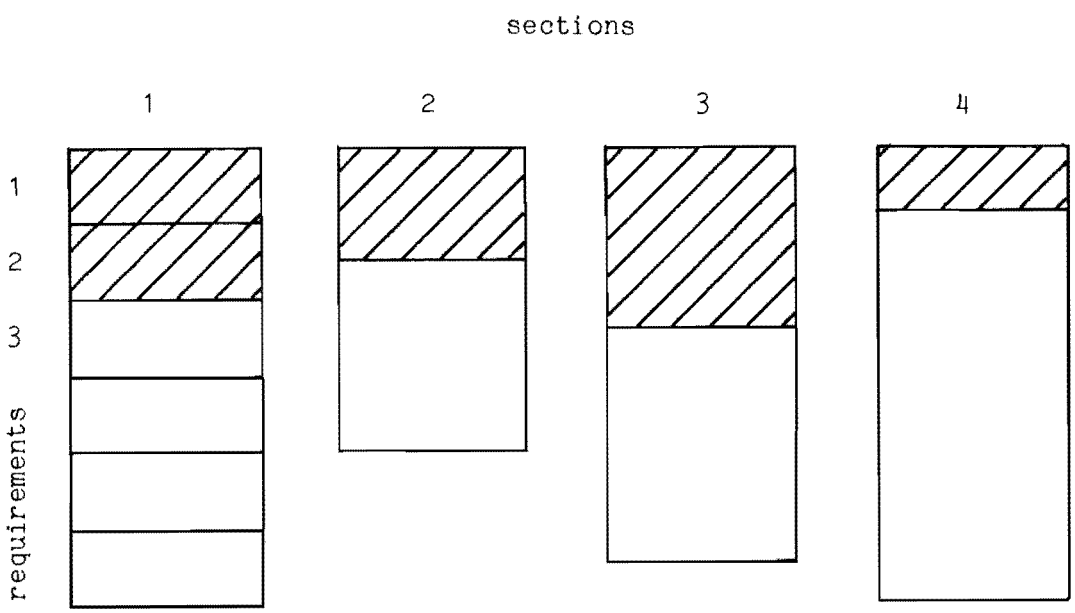


Fig. 3.1 An example of a structured software specification. The specification consists of many requirements. These are divided into independent sections. The requirements in each section are ordered so that the topmost (shown shaded) are the most general. These requirements are implemented as library routines and the remainder by application routines.

'application' routines which call the appropriate library routines. Now, if the specification is altered but the basic requirements remain the same, the library routines need not be altered and changes are confined to the application routines. Also, if only a section of the specification is altered, only the library and application routines corresponding to that section need to be modified.

Identifiers

When the software is coded, the programmer is concerned with manipulating various entities, e.g. a number; a storage location (i.e. a 'variable'); a sequence of storage locations (i.e. an 'array'); an

instruction; a sequence of instructions (i.e. a 'routine'); and 'structured' entities consisting of a group of several other entities. Each entity is referred to by a name, called an 'identifier'. Entities may be related in various ways. For instance, one set of storage locations may be a subset of another, or two sets of instructions may operate on the same data. Such relationships are called 'functional' relationships to distinguish them from relationships between the identifiers of the entities. The latter are called 'symbolic' relationships and usually consist of each of the identifiers having the same group of letters in it. A constant is a simple entity. However, two constants with the same value are not necessarily considered to be functionally related. To be functionally related there has to be a requirement in the specification that they be equal. For instance, if one array of 1024 elements is to be copied to another array of the same size, the sizes of both arrays are said to be functionally related. The second element of the array may have been obtained from an input device and may happen to be equal to 1024. In this case, the value of the second element is not functionally related to the size of the array.

The programmer should select identifiers in such a way as to indicate functional relationships between entities by symbolic relationships between their identifiers. Since modifications to the specifications are expressed, at least in part, by altering the functional relationship between entities, it can reasonably be expected that the software modification is eased if the identifiers are selected in this way. All of the constants in the specification which are equal and functionally related should be given the same identifier. The constant may then be altered to reflect a new specification by changing the definition of the constant's identifier rather than all of the occurrences of the constant itself. A more complicated example is provided by the application of the naming principle to routines. Within the constraints imposed by maximum identifier length and by the need for the identifier to be easily read, routine identifiers should indicate the function of the routine and the nature of the entities manipulated by it. Abbreviations may have to be used because identifiers are usually restricted in length. Finally, a unique common prefix should be used for all identifiers which may be referenced by users of the software. This prevents ambiguities arising with identifiers from other software modules.

3.5 TEST METHODOLOGY

It is implicit in any specification that the software should function correctly. However, only very rarely is the programmer required to prove in a systematic manner that the software works. It is generally accepted that new software contains a few mistakes. When the user comes across a mistake it is reported to the programmer who then corrects it. Eventually, it is hoped, all of the mistakes will be found. This is not satisfactory because of the inconvenience caused to the user. Testing should be as much a part of the implementation as any other facet of the development process.

Unfortunately, complete testing is impossible. To see why this is so, consider the situation in which the the software in theory has no 'internal' states and so can be considered to be an input-output system. The function of such a system may be verified by applying each input state consecutively and checking the output. However, there is no way of ascertaining with complete confidence what the input states of the software are. For instance, a mistake in the software may make an output dependent on the order in which the assumed input states are applied. The mistake has introduced an internal state, so that the software is no longer characterised by the assumed input states. When the software is intended to have internal states , the problem of testing is more difficult still. In the face of this inherent uncertainty as to the nature of the software being tested, a statistical approach must be adopted, even if only notionally. By this it is meant that, even if failure rates are not measured, the programmer's confidence in the correctness of the software is adequate. This confidence should be based upon a test procedure, however.

Testing is expensive. Even trivial examples of software have enormous numbers of input states. So, exhaustive testing is prohibitively protracted. Means must therefore be sought for reducing the number of states applied, yet still retaining a high confidence level.

When a system can be partitioned into subsystems, and each is tested individually, the effort required falls drastically. Instead of being proportional to the product of the number of input states for each subsystem, it is proportional to their sum. In terms of software, each subsystem is a routine. In practice, the number of input states that are applied must be reduced further still. A small number of randomly selected input states could be applied but this might still leave some parts of the routine unexecuted and therefore untested. A more methodical approach is to

incorporate into the test procedure knowledge of the flow of control within the routine. The control structure of the routine is used to classify the input states so that all members of a particular class execute the same parts of the routine. Testing may then be done by using one representation from each class. In this way each part of the routine is executed at least once.

It is apparent that the knowledge of the internal operation of the routine is vital to deriving a sensible test procedure. In fact, if a routine has many internal states it may be quite expensive to test. It is wise therefore to consider test methods early in the development process.

3.6 TEST TOOLS

In this section it is assumed that a suitable test procedure has been devised. Methods for carrying it out are considered.

When hardware is tested the inputs are created by signal generators and the outputs are observed with an oscilloscope or meter, or more sophisticated devices such as 'logic analysers'. The analogues of such equipment do not seem to exist for software on existing commercial systems. At best, one has a 'symbolic debugger' such as VAX-11 DEBUG (Digital 1982a). This is a software tool which gives the user extensive control over the execution of programs written in a high and low level languages. The program may be executed one instruction or one statement at a time. An entity may be referred to by specifying its address or its identifier. The debugger is controlled by 'debugger' commands specified by the user. Debugger commands exist to alter or display the value of a variable or to execute any routine which is a part of the program. One of the disadvantages of a debugger as a software test tool is that it usually manipulates only the simple 'machine' entities supported in hardware by the machine. The 'natural' entities for testing a routine are the entities used for its arguments. Often each of these is a structured entity (¶3.4.2) composed of many machine entities.

One feature of VAX-11 DEBUG, which is particularly useful for test procedures, is that of command files. A test procedure can be translated into debugger commands and these placed in a file. The entire set of commands can then be executed by specifying the appropriate command with the name of the file. However, such command files are difficult to read for several reasons. One is that several commands may be required to manipulate an argument to a routine if it is not a machine entity. Secondly, the

debugger commands are not a structured programming language. For instance, there is no provision to alter the flow of control.

For a number of reasons it is desirable that the test tool and the developed software be closely related. This may be achieved by interfacing the two programs via natural entities and by using a similar command syntax for each program. The case in which the test tool is part of the developed software, so that the software is 'self-testing', is ideal in this respect. By integrating the test tool with the developed software, the total code is reduced because commonly used routines, such as those for input and output, are shared. Also, since the tool is readily available, testing is easier to perform. A test tool which is integrated with the developed software in this manner is referred to as an 'enviroment' module. In ¶3.5 it is explained that efficient test procedures exploit the software structure. Consequently, if the software is modified the test procedures need to be modified as well. This dual modification is facilitated when the test tool and the software are closely related.

A tool designed explicitly for testing rather than for finding mistakes can improve upon a symbolic debugger in several respects. Firstly, it can provide a comprehensive capability for generating test data and for displaying output. Secondly, it can execute test procedures and application routines (¶3.4.2) in a similar manner so that no special knowledge is required to perform a test. Finally, the tool should be able to be extended by the user in as many respects as possible. In particular, it should be possible to define additional entities, so that the natural entities of a particular piece of software can always be used. In Chapter 4, such a tool is described.

3.7 SUMMARY

Traditional software development has several problems associated with it (¶3.3). These are identified and methods for overcoming them are proposed (¶3.4). It is suggested that considerable emphasis be placed upon the way in which a software development task is specified (¶4.2). The possibility that the software will be modified should also be considered when the software is initially specified (¶3.4.1). It is shown that it is impossible to test software completely (¶3.3). A test method is proposed which allows the software to be assumed to be correct with a high level of confidence, however (¶3.5). To implement the method, a test tool which allows access to the intermediate results

generated by the software is required (¶3.6). Currently available tools such as symbolic debuggers are considered and found to be insufficient. It is suggested that a tool which is integrated with the software itself has additional benefits. Such a tool is here called an 'enviroment' module . In Chapter 4 an enviroment module suitable for developing image processing software is described.

PART II

NEW RESULTS

4. A SOFTWARE MODULE FOR IMAGE PROCESSING

4.1 INTRODUCTION

In this chapter a software module to facilitate algorithm development and programming testing is described. The module is particularly suited to an image processing research enviroment. The design philosophy is presented in ¶4.2 and an overview of the module is given ¶4.3. The user controls the module by supplying commands either interactively from a terminal or indirectly from a file. The command syntax is introduced in ¶4.4 and the types of data which the user can manipulate are detailed in ¶4.5. The interactive nature of the module is discussed in ¶4.6 and input from files is covered in ¶4.7. One of the special features of the command syntax is that the available commands are subdivided into groups called 'command modules' and this is discussed in ¶4.8. Commands for implementing program loops and for converting data from one form to another are presented in ¶¶4.9 and 4.10 respectively.

4.2 DESIGN PHILOSOPHY

The GET Enviroment Module (GEM) was developed in response to the need outlined in Chapter 3 for an enviroment module (¶3.6) that supports testing and encourages software sharing.

There are two ways in which the enviroment module and the software ('user module') that the user of the enviroment module has written may be related. In the method illustrated in Fig. 4.1(a), the user module is executed or 'dispatched' from the enviroment module. This approach is used by the Command Definition Utility described by Digital (1982b). A difficulty with this method is that the user module must have a fixed calling format. This is because the enviroment module has to be able to call a variety of user modules and so a simple fixed interface is necessary. The alternative is for the user software to call the enviroment module to <u>get</u> control information and data. This is depicted in Fig. 4.1(b). Well defined interfaces are still used but now the user module can have a more arbitrary form. The GEM may be expanded by incorporating routines from user modules, which use the standard interface, into it. The extra capability of the extended GEM is then available to other user modules.

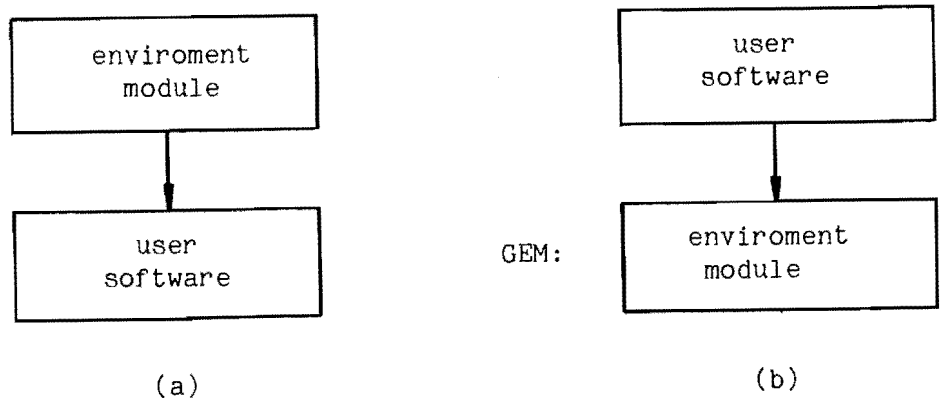


Fig. 4.1 Design alternatives for an enviroment module. In (a) the enviroment module calls the user software, while in (b) the user software calls the enviroment module for control information and data.

4.3 OVERVIEW

In this section an outline of the usage of the GEM is presented. The software is run on a computer with VAX-11 architecture (Digital 1979), running the VAX/VMS operating system (Digital 1978). Software is developed with this operating system by creating a 'source' file with an text editor, compiling the source into an intermediate 'object' code and then 'linking' the object code with other object files and with the results of previous link operations. The result of the link is an 'image' file which may be executed.

When the image is executed, the user controls the software by supplying command lines. This input may originate from a terminal or from a file. Terminal input is processed by the GEM interactively. The user is prompted for input and is allowed to take corrective action when an error occurs. Typically, each command causes a particular routine to be called. Consequently, there is usually a separate command for each library routine (¶3.2), and an additional one for the application routine (¶3.2). Thus, the library routines may be tested and the application routine executed by entering commands with similar syntax. The entities to be manipulated by the routine invoked by the command are selected by supplying their

identifiers (¶3.4.2) as 'parameters' to the command. Entities such as scalars, vectors, matrices, and images may be processed by the GEM.

The detection of mistakes is facilitated by having one command for each routine. Since each routine can be executed independently, a mistake in one does not prevent the rest from being tested. This is in contrast to the traditional method of calling one routine after another from a main program. If a 'fatal exception' occurs the program 'crashes' and the routines called after the incorrect routine cannot be tested without fixing the mistake which caused the fatal exception.

Routines to perform input and output are an essential part of any enviroment module. The GEM supports two types of input/output routines: 'print' and 'move'. Print routines are used by the GEM to process the command line and to output the result in a format which can be read by a person. Move routines transfer data between memory and mass storage devices. The format of the data is not altered by a move routine and therefore the transfer is quicker and the result more compact than for a print routine. Several transfers may be in progress at the same time. A 'stream' entity controls each transfer.

Test procedures are implemented as command files and are executed by referring to the name of the file. Documentation for the test procedures may appear in the command file since comments are permitted anywhere in the syntax.

Prior to using the GEM, the user must acquire a minimal understanding of the syntax and terminology employed by the GEM. Later sections of this chapter supply the requisite knowledge. During execution, information can be provided by the software as to the syntax of all of the commands which are available. When used interactively, the GEM supplements this by prompting the user for all expected but unspecified input, indicating defaults and requirements. If the input supplied by the user is inappropriate for some reason, the user is given the opportunity to enter a correct version.

In addition to the programming required for the library and application routines, a certain amount of extra programming is required to interface these routines to the GEM. Fig. 4.2 indicates the resulting structure. In addition to the library routines and application routines, the user supplies a 'parse' routine to control all of the other routines and to transfer data between the GEM and the library and application routines. The parse routine calls the GEM to obtain commands and

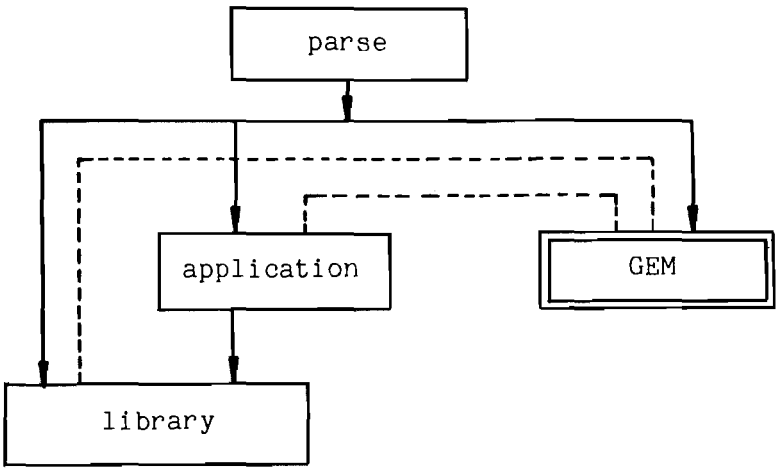


Fig. 4.2 The interface between the user module and the GEM. A box with a single outline represents software written by the user. A solid arrow from one box to another indicates that routines associated with the first box call routines represented by the second box. The dashed lines indicate the flow of data.

parameters. It then decodes the commands and calls the appropriate user routines and GEM output routines, if these are required. The structure of the parse routine is well defined and it would normally be created by modifying an existing parse routine.

4.4 COMMAND STRUCTURE

The format of a complete GEM command is

'C P1 P2 ... PN',

where the command format is the text enclosed by the quotes, the ellipsis denotes repetition as necessary, C is the command 'keyword', and P1 to PN are the command parameters. The command elements are separated by an arbitrary, nonempty juxtaposition of blanks and tab characters. In some cases, when the parameters consists of special characters, explicit 'separators' may not

be required.

The keyword may be entered in a mixture of upper and lower case letters and may be abbreviated, provided that the shortened form unambiguously distinguishes it from other commands. Throughout this thesis keywords are shown with their initial character capitalised. A 'Help' command displays a list of the commands available.

Each parameter belongs to a particular class (for instance, scalars or matrices), and is either a variable or a constant. A variable is identified by specifying a symbol and a constant is denoted by a 'literal' describing the particular value intended. When a literal is specified for a variable parameter, a variable is created for the duration of the command with the literal's value.

Arbitrary numbers of 'qualifiers' may appear after either the keyword or each parameter. The syntax of a qualifier is determined by the parse routine, but the initial character must be special and is typically a '/', '+' or '-'.

A complete command may be entered without specifying all of the parameters by terminating the command with the end-of-command character, which is denoted by <eoc>. The <eoc> character may be altered by the user but is usually set to the semicolon. Any number of commands may be supplied in a single command line, provided that they are terminated either by specifying all of the parameters expected by the parse routine or by the <eoc> character. Commands read from a file are automatically terminated by the end of the record, unless the last character is the continuation character (denoted by <cont>), which is usually a hyphen.

Comments may appear only at the right of command lines and are begun with the comment character (denoted by <comment>), which is usually the exclamation mark.

4.5 PARAMETER CLASSES

The VAX-11 calling standard (Digital 1982b Appendix C) defines the classes that an argument to a routine may have. A subset of these, consisting of the more commonly used ones, is currently implemented by the GEM.

In response to a request from the parse routine for a parameter, the GEM begins parsing the command line from left to right, starting from the end of the previous parameter or command keyword. Blanks and tabs are

Table 4.1 GEM basic tokens. The description specifies the characters which are acceptable in the token.

Class Name	Description of Token	Examples
Any	Next non-blank, non-tab character.	'+', 'A'.
String	All upper and lower case alphanumeric characters, underscore and dollar sign.	'abAB$_'.
Key	All string characters, and the token is converted to uppercase.	'a' and 'A' are equivalent.
Filename	Any characters valid in an RMS file specification, i.e. string characters and any of the characters in the string '[]<>:;.%*+-'.	'USER:[MINARD.GET]GET', '*;-3', 'SYS$LIBRARY'.
Number	Any of the characters in the string '+-0123456789.Ee'.	'1', '-1.', '13', '1E7', '1e-7'.

skipped. Each parameter is specified by a 'token' which is a sequence of characters. For each request only one class is expected, and this is used to determine which characters are accepted as a constituting a valid token. All of the tokens can be constructed by juxtaposition or by additional semantics from five 'basic' tokens. Each of the basic tokens is capable of specifying a member of a parameter class by itself. These classes are called 'basic' classes and are enumerated in Table 4.1. The remainder of the parameter classes are referred to as derived classes. A filename token may not in fact have a valid file name syntax as defined by the Record Management System (RMS) (Digital 1982d) but this is detected when an attempt is made to open the file. The additional semantics used to construct the derived classes are implemented by extra processing in the GEM after parsing the basic token(s). Table 4.2 lists the derived classes in approximately increasing order of complexity. The numeric conversion

required to obtain an 'integer' value from the string of characters which constitute a 'number' token is an example of the extra processing required to obtain a derived class from a basic one.

Table 4.2 GEM derived tokens. The description specifies the syntax of the token using the notation: [] to enclose an optional item; {} to indicate that a choice of one of the enclosed items is required and []... to denote repetition as necessary. The examples are enclosed by quotes.

Class Name	Description of Token and Associated Processing	Examples
Integer	A number which is converted to integer^4.	'12'.
Real	A number which is converted to real^4.	'12', '12E-13'.
Text	A string or any sequence of characters enclosed in double quotes.	'a', '"A+B"'.
Code	A key which is looked up in a code table to produce an integer^4 value. See Table 4.3 for some of tables available. Table lookup allows ambiguous matches, by accepting the first valid abbreviation found. This class provides a named constant facility.	
Data_type	data_type code[^length]	'real^4', 'integer'.
Symbol	A key which is looked up in a symbol table to yield an integer^4 value. Abbreviations must be non-ambiguous. The symbol table may be altered and the values are usually interpreted as an address to provide variables.	

Table 4.2 (cont.)

Class Name	Description of Token and Associated Processing	Examples
Field	A symbol which refers to a data structure specifying the origin and size of a rectangular region in two dimensions.	
Print	A symbol which refers to a data structure, called a 'print block', specifying the attributes of a human readable output stream. These attributes are the maximum record width, the maximum width written, the number of lines written and the indentation from the left margin to be used for further output.	
Move	A symbol referring to an unformatted input/output stream.	
Area	A symbol referring to a text storage area. Such an area has a variable width and length.	
Scalar	A symbol or literal of the form: (data_type value), which specifies a numeric value of a certain data_type.	'A', '(complex^8 2 1)'.
Vector	A symbol or literal of the form: (data_type [value]...)	'A', '(real 3 2 1)'. The latter is referred to as a real^4 3-vector.
Array	A symbol referring to an array of numeric values. Associated attributes are the number of dimensions, the bounds of each dimension and the data_type of the values stored in the array.	

Table 4.2 (cont.)

Class Name	Description of Token and Associated Processing	Examples
Matrix	An array with two dimensions, each with a lower bound equal to 1. A symbol or literal of the form: (data_type [[value].../]...)	'A', '(r 1 2/3 4/5 6)'. The latter corresponds to 1 2 (3 4) 5 6 which is an example of a real^4 2×3 matrix.
Image	A matrix and a field, specified by 'matrix+Field=field' or 'matrix'. In the latter case the field symbol is assumed to be the same as the matrix symbol. The field specifies a coordinate system for the image.	

Table 4.3 GEM code class tables. The 'function' code class provides only generic function names. The actual function used is determined in context by the data types of the domain and range. The 'facility' codes provide names for each of the parameter classes.

Table Name	Examples
Horizontal Justification	'left', 'centre', 'right'.
Vertical Justification	'bottom', 'middle', 'top'.
Data_type	'unsigned', 'integer', 'real', 'complex', 'character, and all VMS defined data types (see Digital 1982c Appendix C).
Function	'sin', 'cos', 'tan'.
Facility	'image', 'scalar'.

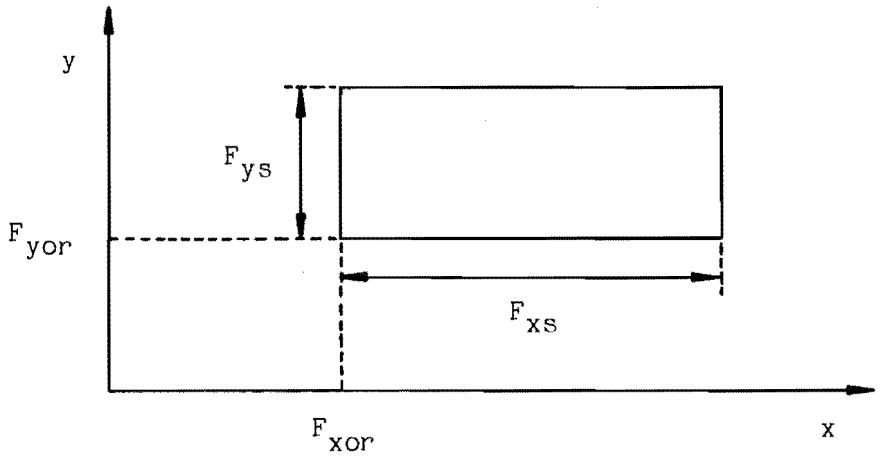


Fig. 4.3 Field parameter class definition.

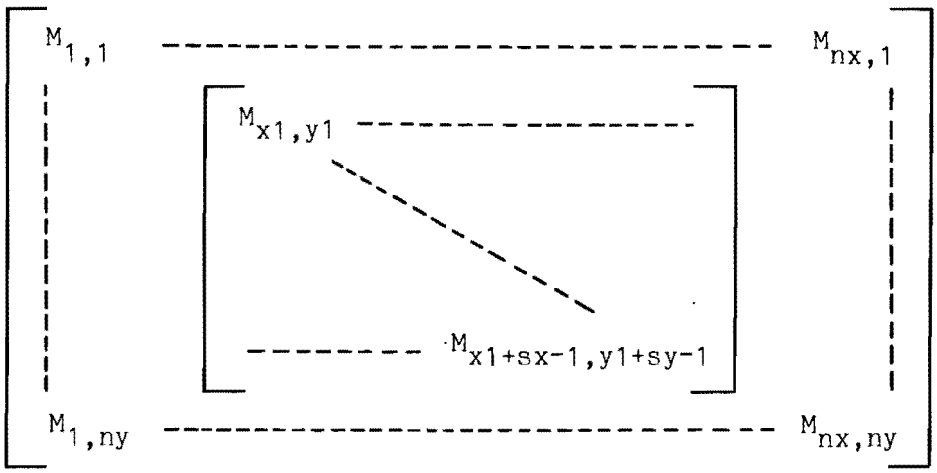


Fig. 4.4 Matrix parameter class definition.

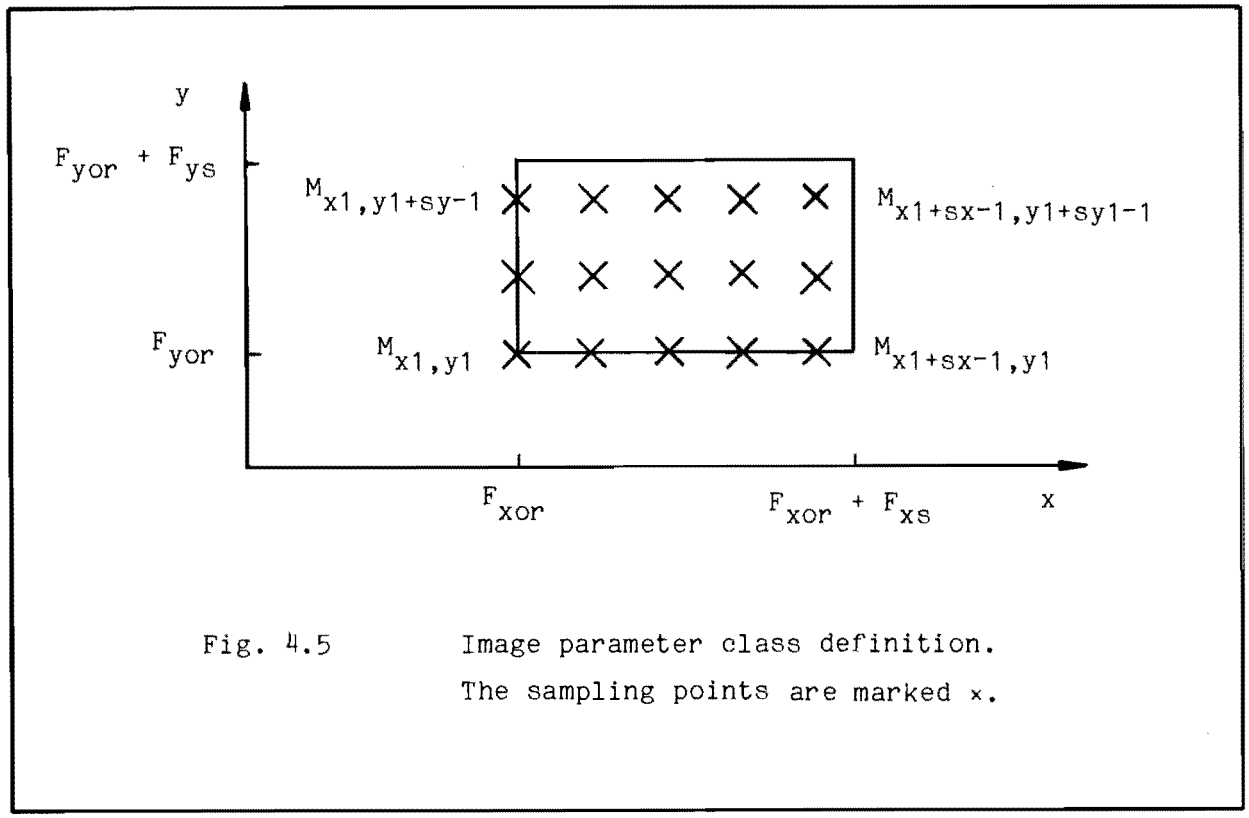


Fig. 4.5 Image parameter class definition.
The sampling points are marked ×.

4.5.1 Fields, Matrices and Images.

Image processing software using the GEM has been developed to implement the ideas in this thesis. The parameter classes fields, matrices, and images are widely used in this software and are therefore described in some detail here.

A 'field' describes a rectangular region in two dimensions specified by four real^4 values (a 'real^4' value is a representation of a real number stored in four bytes), denoted by F_{xor}, F_{yor}, F_{xs} and F_{ys}, which are defined in Fig. 4.3.

A 'matrix' is a two-dimensional array of values embedded in a larger two-dimensional 'parent' array $M_{i,j}$. Six parameters sx, sy, nx, ny, x1, y1 describe the embedding. The parent array is of size nx × ny, the smaller 'subarray' is of size sx × sy. The origin of the subarray is at $M_{x1,y1}$. This is illustrated in Fig. 4.4. Note that the display convention is the transpose of that adopted in mathematics.

An 'image' is the combination of a matrix and a field. The values in the matrix are considered to be samples of a function taken on a regular grid defined by the field. The sampling pattern adopted is shown in Fig. 4.5 and the relationship between the coordinates (x,y) and the matrix subscript (i,j) for the same sample is

$$x = (i - x1)\ F_{xs}/sx + F_{xor} \quad , \tag{4.1}$$

and

$$y = (i - y1)\ F_{ys}/sy + F_{yor} \quad . \tag{4.2}$$

This particular sampling grid was chosen because it has two useful properties. First, two such grids may be juxtaposed without introducing irregularities at the boundaries of the field, and second, the origin of the field is sampled exactly.

4.6 THE GET OPERATION

Data is passed between the GEM and the parse routine by calling a 'get' routine within the GEM. Each of these calls gives rise to a 'get' operation, which consists of translating unprocessed characters in the command line into data which the parse routine then passes on to the user software.

The parse routine specifies which class of parameter it requires by calling the appropriate get routine. However, the parse routine may require only members of the class with particular attributes. The additional requirements are specified by arguments to the get routine. The attributes concerned might be the number of dimensions in an array, the size of a a particular dimension, or the data type of an array. If the parameter parsed does not satisfy the requirements a standard error recovery procedure is invoked.

The recovery procedure begins with display of an error message. Then the rest of the command line is discarded. If the GEM is being used interactively, the get operation is repeated. If not, an error status is returned to the parse routine, which then stops requesting parameters and requests a new command.

A default response may be supplied by the parse routine for each get operation. This response is assumed if the user enters an empty line (i.e. just presses the return key), or if the next token is the default character ('#'), or if the end of the command line is reached.

An interactive get operation can be aborted by typing control-Z. The parse routine responds to this by skipping to the next command.

When a parameter is requested and there are no tokens left in the command line, the action taken depends on whether the get operation is interactive or not. In a file, an <eoc> condition (¶4.4) is generated, which causes the default to used. An alternative option is available if the GEM is being used interactively and provided an argument identifying the parameter has been specified. This argument is a string and is used to construct a prompt which is displayed, indicating that more input is required. Prompts for command keywords end in '>', for parameters in ':', and for qualifiers in '?'. The initial character of a qualifier prompt indicates the initial character to be used for the qualifier. Typically, this is '/', but '+' and '-' are also used. Information on requirements may appear within square brackets and the default, if present and not blank, is shown between parentheses. Examples of the three types of prompts are shown in Fig. 4.6.

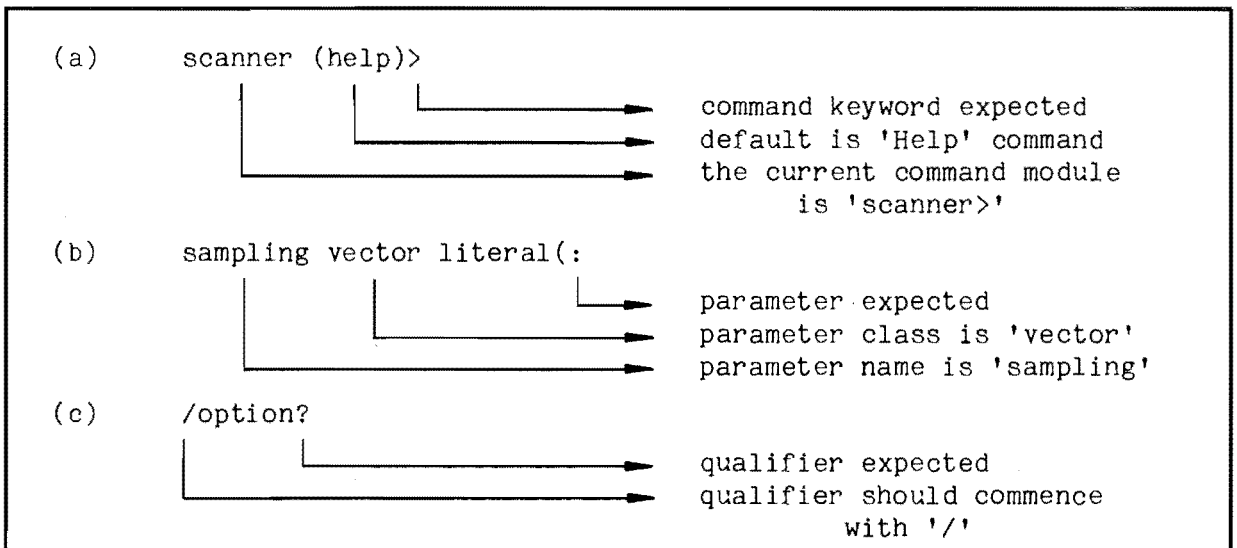


Fig. 4.6 Prompt formats and their interpretation.
(a) Command. (b) Parameter. (c) Qualifier.

4.7 COMMAND FILES

Command files allow sequences of commands to be stored. The file is entered (i.e. the commands in the file are executed) by the 'indirect file'

command

'@filename',

which directs the input stream to the file. The indirect file command also terminates any previous command and so can cause defaults to be used for the previous command. The file is exited when an end-of-file (denoted by <eof>) condition occurs. This condition arises when the end of the file is encountered or when the command 'Eof' is parsed. When the file is exited, the file is closed and input reverts to the terminal or the previous file.

An interactive user may cause the remainder of the commands in a file to be ignored by typing control-C. This causes an <eof> condition to be present until the input stream is derived from the terminal.

4.8 COMMAND MODULES

The set of all commands available in the GEM are partitioned in a logical manner into subsets. Each subset is implemented by a 'command' module. At any one time, only the commands implemented in the 'current' command module are available. There is provision, however, for changing the current command module, not only in between commands, but in the midst of them as well. Command modules can therefore be nested.

This modular approach benefits both the programmer and the user. For instance, all of the commands for an application may be placed in one module and implemented without affecting the command structure of other modules. Hence, the programming task is easier. The output from the Help command is uncluttered with commands of other, possibly irrelevant, modules because related commands appear in the same module.

Each module has its own parse routine and is identified by some easily typed name. To enter a module called 'get', the user types:

'get>'.

At any particular time, a default module exists and may be entered by typing

'>'.

When a command keyword is expected, the default module is 'get'; when a parameter is expected, the default module is usually capable of manipulating entities of the same class as the parameter. Typing,

'?>'

displays a list of modules available in the current context and indicates

the default module.

A command module is exited by typing:

'Exit', 'Quit', or control-Z.

The latter is interpreted as for the Exit command. When the image exits all of the data maintained in memory by the image is lost. To help prevent this occurring accidently, the Quit command must be used to exit from the outermost module.

Often it is desired to enter a module, execute a few commands, and exit. A special syntax for the module entry command facilitates this: when module entry is followed by a command on the same line, a module exit is generated at the end of the line. The net result is that the current module at the beginning of the next line is the same as that for the beginning of the current line. In a similar manner, module exits are generated at file exit (¶4.7) to restore the current module to the one current at file entry.

For most parameter classes there is a corresponding module which manipulates entities of that class. For the classes scalar, vector, and matrix, the commands are further divided into three modules according to their function. The most basic module (scalar>, vector>, and matrix> respectively) performs descriptive tasks, another module (sio>, vio>, and mio>) performs input/output and the third module (sop>, vop>, and mop>) performs operations. For example, the descriptive module has a 'Declare' command to create an entity with a given identifier and attributes, a 'Format' command to display the attributes, and a 'Release' command to delete an entity.

Lists, called 'symbol tables', of all the entities that have been defined are maintained. These lists are declared by the user module and passed to the GEM as a routine argument. A symbol> command module exists to manipulate the symbol tables.

A module such as symbol> needs access to only one table at a time. It would be possible to expand the symbol> module to include all of the tables currently present in the GEM. This is not done because it would increase the complexity of the symbol> module and more importantly it would complicate extensions to the GEM. To see this, consider the changes which would need to be made if a command module were assimilated into the GEM which used another symbol table. The new table would have to be added to the argument list of the symbol> module and to every module that allows the

symbol> module to be called from it, and so on. The set of symbol tables accessible from a command module is called the 'data context' of a module. In general, the smaller the data context of a command module, the simpler the module is to code and use.

Some of the parameter classes belong to a natural hierarchy which results from the way they are manipulated and defined. For instance, the data context of the vector> module must include the scalar symbol table if it is to declare a vector of a length given by a particular scalar. There is no analogous requirement in the scalar> module to 'know' about vectors. For this reason the vector> module need not be callable from the scalar> module.

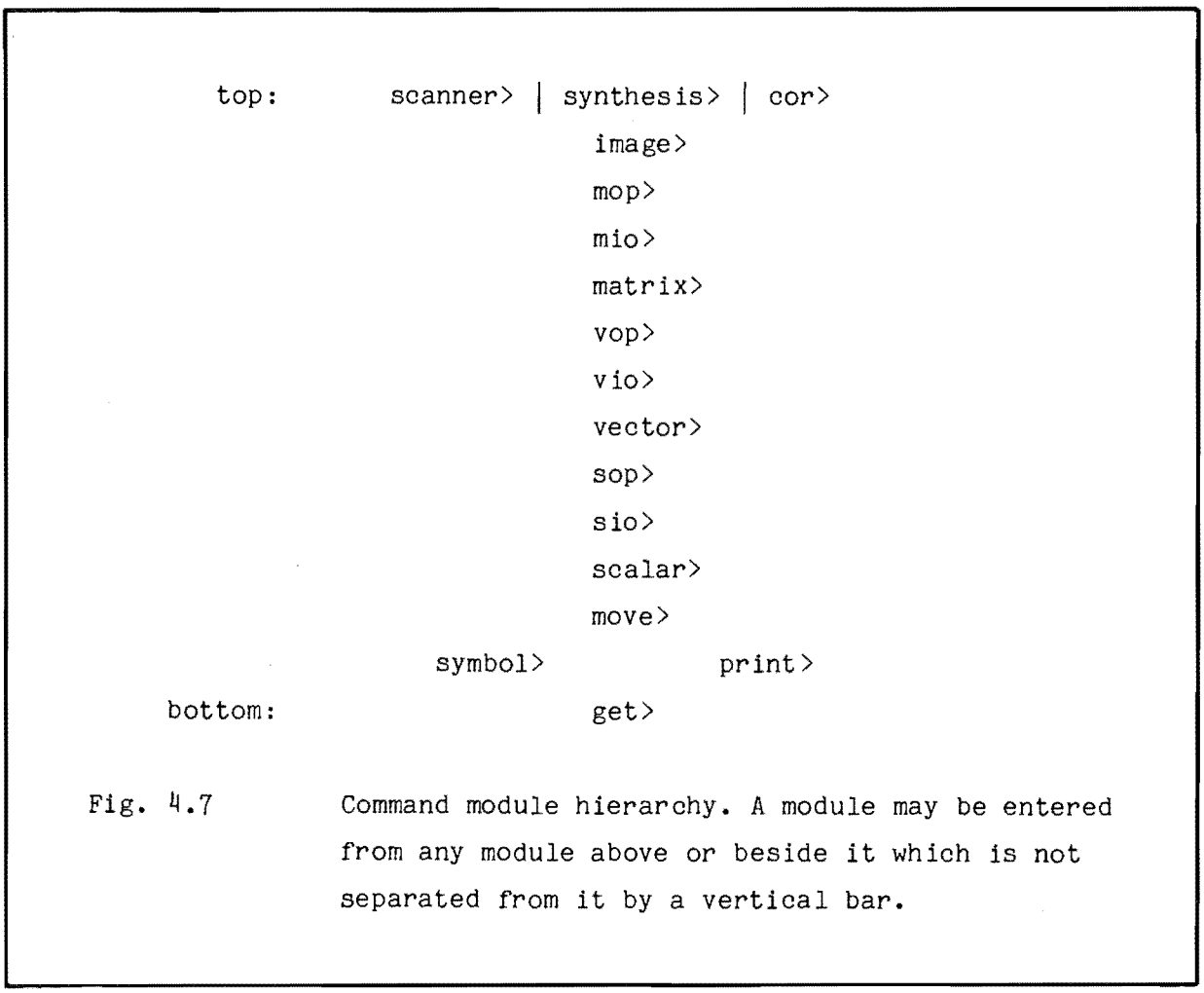


Fig. 4.7 Command module hierarchy. A module may be entered from any module above or beside it which is not separated from it by a vertical bar.

To eliminate the difficult modifications that are associated with using large data contexts, and to take advantage of the small data contexts of some modules, a hierarchy is imposed on the command modules. This is shown in Fig. 4.7. The hierarchy is imposed by the rule that a module may be

entered from any module above or on the same level as itself, providing that a recursive calling sequence does not result.

The commands available in the image>, mop>, mio>, matrix>, and field> modules are listed in Appendix C.

4.9 COMMAND LOOPS

A sequence of commands may be repeated with the aid of 'loop' clauses. These clauses are defined as shown in Fig. 4.8. In the figure, keywords are shown in uppercase, 'block' represents a sequence of commands in which the final command is terminated, 'l_sca' and 'i_sca' are logical and integer scalars respectively. A loop is said to be 'active' when the commands belonging to it are being executed and the commands are not part of a loop nested within the first loop. Any active loop may be terminated with the End clause. This is interpreted as End_Do, End_Repeat or End_While according to the active loop.

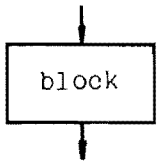
The semantics of the loop clauses are shown by the flow diagrams in Fig. 4.8. The block of a Do clause is executed at least once, in contrast to the While clause in which the block may not be executed at all. If the integer scalar of a repeat command is zero or negative, the block is not executed.

Command loops may be nested arbitrarily, but the beginning and end clauses must occur in the same command module. This is because each clause may refer to the data context, and this must be guaranteed to be the same at loop exit and loop entry. If a command module exits, and there are loops still being executed which were entered in that module, End clauses are generated until all of the loops are terminated.

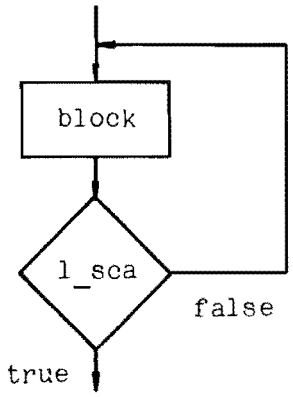
4.10 SLICING AND CLASS CONVERSION

It is possible to define a new entity consisting of a subset of values from an array. When the subset is specified by restricting the range for each subscript, the resulting entity is known as a 'slice' or a 'noncontiguous' array. The latter terminology is suggested by the storage locations of the new entity not being contiguous. An array is said to be stored by row if the storage locations of all of the elements in a row are contiguous.

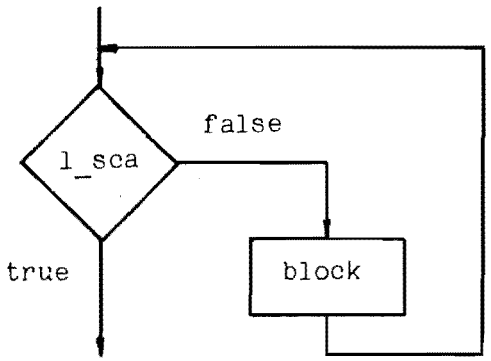
Do block End_Do



Do block Until l_sca



While l_sca block End_While



Repeat i_sca block End_Repeat

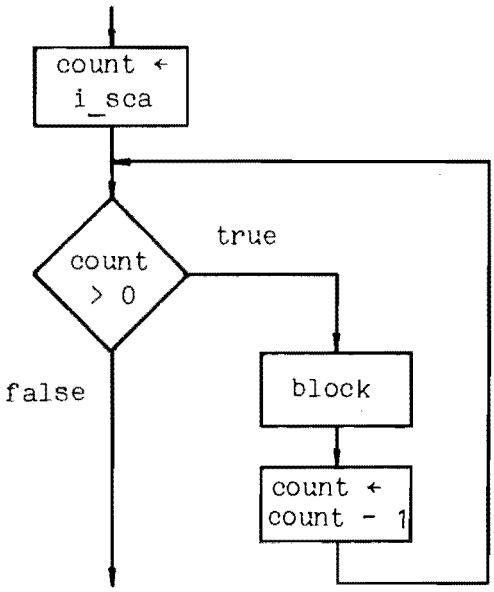


Fig. 4.8 Loop clauses.

Some parameter classes can be regarded as special slices of others. For instance, a scalar is a slice of one element of a vector, a vector is a slice of one row of a matrix stored by row. A vector is not considered to be a slice of one column of a matrix stored by row because the storage locations corresponding to the column might not be contiguous.

```
matrix_option   = { Sub_Field = sub_field,
                    {Point,XY} = point,
                     X = x,
                     Y = y,
                     X{bounds,range} = x_bds,
                     Y{bounds,range} = y_bds,
                     Help}
image           = matrix[+matrix_option]...
```

Fig. 4.9 Image slice syntax. The conventions used to specify the syntax are described in ¶5.4.3. 'sub_field' is a field. 'point', 'x_bds', and 'y_bds' are real^4 2-vectors. 'x' and 'y' are real^4 scalars.

The GEM provides a method for converting one class to another and for 'slicing' (i.e. obtaining a slice from an entity). The syntax for specifying a parameter includes additional syntax to specify a slice. Additional qualifiers are used to specify an image slice and these are shown in Fig. 4.9. An image slice may consist of a row (+Y qualifier), a column (+X), a pixel (+Point), or a range in one or both of the subscripts (+Sub_Field, +Xrange, +Yrange). The restrictions on the x and y subscripts are handled independently, and the last specification for each is used to form the slice. If no restriction is specified the entire range of the coordinate is used.

Class conversion is provided by defining a class called 'facility_matrices', which also includes provision for slicing matrices and vectors. The syntax is shown in Fig. 4.10.

```
matrix_option   = { {Point,XY} = point,
                    X = x,
                    Y = y,
                    X{bounds,range} = x_bds,
                    Y{bounds,range} = y_bds,
                    Help}

vector_option   = { Element = e,
                    Bounds = e_bds}

facility_matrix = { Image image,
                    Matrix[/matrix_option]... matrix,
                    Vector[/vector_option]... vector,
                    Scalar scalar
                    facility/Help}
```

Fig. 4.10 Facility_matrix syntax. 'point', 'x_bds', 'y_bds' and 'e_bds' are integer^4 2-vectors. 'x' and 'y' are integer^4 scalars.

```
value_matrix    = { matrix,
                    values> facility_matrix}
```

Fig. 4.11 Value_matrix syntax.

A 'value_matrix' class is constructed from matrices and facility_matrices to allow the latter to be supplied as a parameter to commands which usually have matrix parameters. The syntax of this class is

shown in Fig. 4.11.

Two other classes, 'facility_vectors' and 'value_vectors' are defined in a similar manner to facility_matrices and value_matrices. However, images and matrices may not be sliced to form facility_vectors because the data context (¶4.8) associated with facility_vectors does not include matrices.

4.11 SUMMARY

A software module, called the 'GET Enviroment Module (GEM)', is described in this chapter. The module is designed to facilitate the development of image processing software in particular, but is useful in general. A standard interface is provided between the user and the program. The interface is defined so that interactive and batch usage is very similar. Several entities are defined by the GEM including three of particular importance in image processing. One of these stores a sampled representation (¶1.2.5) of an image.

The GEM module is extended in a well defined manner to implement a particular software specification. The software written by the user to implement this extension is called a 'command module'. Several command modules have been written by the author to implement methods proposed in this thesis. A 'scanner>' command module is described in ¶5.5 which controls the ultrasonic scanner used to obtain the experimental results presented in Chapters 6 and 7. In ¶7.3 a 'cor>' command module is described which implements an extension of a speckle processing algorithm called 'shift-and-add'. A 'synthesis>' command module described in ¶7.6.1 was used to generate the objects for the simulations reported in this thesis.

5. AN ULTRASONIC SCANNER - DIGITAL COMPUTER INTERFACE

5.1 INTRODUCTION

This chapter describes the interfacing of an ultrasonic scanner to the VAX 11/750 minicomputer (VAX) in the Computer Laboratory of the Electrical and Electronic Engineering Department of the University of Canterbury. The scanner, built by Dunlop (1978), was originally interfaced to an EAI-590 hybrid computer. However, in December 1983 the digital half of the hybrid computer was scrapped. So, in order to allow the experimental work (reported in this thesis) to proceed, the scanner had to be interfaced to another computer and the VAX was chosen for this purpose.

The scanner is described in ¶5.2 and the signals required to use it are discussed in ¶5.3. The hardware and software for the interface are described in ¶¶5.4 and 5.5 respectively. The calibration procedures performed to complete the interface are explained in ¶5.6. Finally, in ¶5.7 the operation of the scanner and VAX as a CAI system (¶1.5.6) is discussed.

5.2 A DESCRIPTION OF THE ULTRASONIC SCANNER

Several descriptions of the ultrasonic scanner with varying degrees of emphasis have been written. Dunlop (1978 Ch. 4) describes the scanner as a system for obtaining projection measurements for CT (¶1.7.1). Robinson and Bates (1980) and Robinson (1982) explain how the scanner, with slight modifications, is used to obtain measurements of ultrasonic diffraction patterns. Nhunduru (1984 Ch. 3) treats the scanner as a general ultrasonic measurement facility. The description presented here assumes that the scanner is part of a CAI system (¶1.5.6) and is based on the general theory of such systems as presented in Chapter 1.

The ultrasonic scanner is a device for making ultrasonic scattering measurements and as such may be combined with a digital computer to form a CAI system as described in ¶1.5.6. The physical system (outlined in ¶1.5.1) in which the ultrasound propagates consists of a reference medium, a generation region, a measurment region, and an object region. The realisation of each of these regions in the scanner is described in ¶¶5.2.1, 5.2.2 and 5.2.3 respectively. In the scanner the generating and measuring regions may be moved relative to the object region. The way

in which this is achieved is explained in ¶5.2.4. Amplitude control of the generating sources and measurement of the receiving sources is described in ¶5.2.5.

5.2.1 The Reference Medium

The reference medium is a cylindrical body of water partially filling a galvanised iron tank (Fig. 5.1). Two precautions are taken to ensure that the reference medium is homogeneous. The first is that the water is allowed to come into thermal equilibrium with the laboratory which is maintained at a temperature of 23°C. This precaution is taken because the velocity of sound in water is temperature dependent. At the temperature of the laboratory, it is (Wells 1977 ¶1.3)

$$c_0 = 1495 \text{ m/s} . \qquad (5.1)$$

The second precaution is to dislodge any air bubbles that may be attached to the tank wall or to other parts of the apparatus immersed in the water.

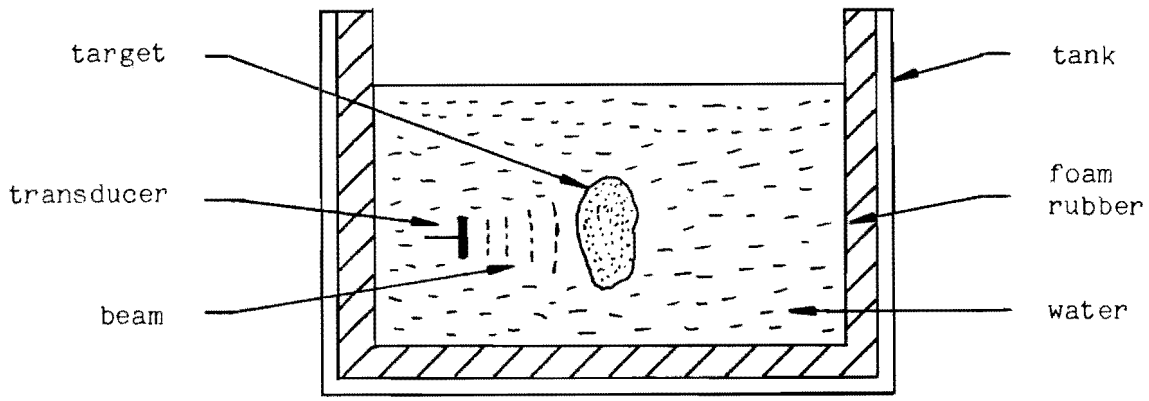


Fig. 5.1 A cross-section of the ultrasonic scanner tank. The tank, which contains the target and the transducers, is filled with water. The tank walls are lined with foam rubber to minimise reverberation.

There are obvious conceptual advantages in treating the reference medium as an unbounded propagation medium. The water may be treated as an unbounded medium if its boundaries do not affect the scattering measurements. This is (approximately) achieved by a combination of several methods. The wall and base of the tank are lined with foam rubber. This absorbs ultrasound so that the amplitude of any incoming waves resulting from the reflection of outgoing waves is negligible (Robinson 1982 ¶5.3).

The water-air interface is effectively a perfectly reflecting boundary, however, since the acoustic impedance (¶1.5.7) of water (Wells 1977 ¶1.7)

$$Z_o = 1.5 \times 10^{-6} \text{ kg m}^{-2}\text{s}^{-1} \tag{5.2}$$

and air (Wells 1977 ¶1.7)

$$Z_{air} = 4 \times 10^{-10} \text{ kg m}^{-2}\text{s}^{-1} \tag{5.3}$$

are so different. In fact, the reflection coefficient $R_{0,air}$ (Wells 1977 ¶1.8) is

$$R_{0,air} = \frac{Z_{air} - Z_o}{Z_{air} + Z_o} \tag{5.4}$$

$$\approx -1 \ .$$

Reflections from the interface are minimised by ensuring that the main lobe of any transducer immersed in the water does not impinge upon the surface.

The Green's function for the reference medium can therefore be taken as (1.183) or (1.126) for two or three-dimensional scattering geometries respectively.

5.2.2 The Generation and Measurement Regions

Piezoelectric transducers (¶1.5.7) of similar construction are used for both generation and measurement. The transducers were fabricated in the Electrical and Electronic Engineering Department's workshops and their construction is described by Dunlop (1978 Ch. 5). Backing and matching sections are used to couple the acoustical energy into a water load over a wide range of frequencies. The maximum response of the transducer is at 2.5 MHz and the half-power bandwidth is 1.2 MHz (Dunlop 1978 ¶5.6). The piezoelectric element is a 10 mm diameter Phillips PXE5 PZT disc, with a beam width of (Dunlop 1978 ¶5.6.2)

$$\theta_{BW} = 12^\circ \pm 3^\circ \tag{5.5}$$

at 2 MHz.

The arrangement of the transducers is shown in Fig. 5.2. Ultrasound is generated by the transmitting transducer T and measured by the reference path transducer R and the scattered path transducer S. The correspondence between Fig. 5.2 for the scanner geometry and Fig. 1.5 for the general remote probing geometry is established by noting that the unit wave vector $\hat{\mathbf{k}}$ for the incident wave is

$$\hat{\mathbf{k}} = \hat{\mathbf{x}} \tag{5.6}$$

so that the rotated coordinates (ξ,η) are given by

$$\xi = -y \tag{5.7}$$

and

$$\eta = x \ . \tag{5.8}$$

Projection measurements (¶1.7.1) are obtained from transducer R and far-field measurements (¶1.7.2) from S. When S is used for far-field measurements, the far-field condition (1.216) places a restriction on the maximum size of the object region.

5.2.3 Object Region

In this subsection, various methods for supporting objects within the object region and for creating inhomogeneities in the object and reference regions are presented.

Robinson (1982 p.128) describes a method suitable for supporting objects consisting of several parallel strands. The method is depicted in Fig. 5.3. Two perspex plates with a hole drilled for each strand of the object are placed one above and one below the object region. The strands are threaded through the holes and clamped to the top plate. A weight is attached to the bottom of each strand to keep it taut. The suspension system is suitable for arbitrary arrangements of a small number of strands. It is not suitable for large numbers of strands because of the time required to thread and weight each strand.

To overcome this difficulty a 'grid object holder' was constructed. This holder may be easily removed and replaced from the object region and is capable of suspending a linear array of strands. As shown in Fig. 5.4 the holder consists of four parallel brass 5 mm rods which are held by a U

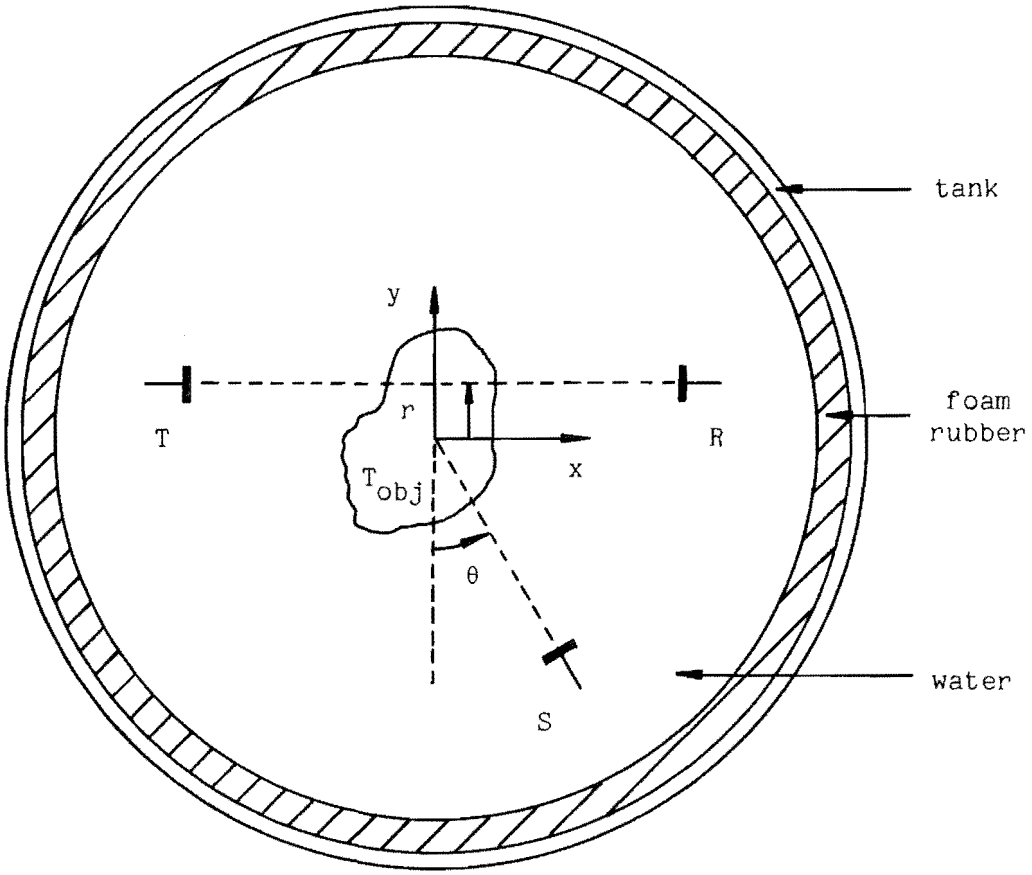


Fig. 5.2 A horizontal cross-section of the ultrasonic scanner. The generating region T_{gen} is occupied by the transmitting transducer T. The receiving region T_{meas} contains the reference path transducer R and the scattered path transducer S. The coordinates (x,y) of T, R and S are (-13cm,r), (13cm,r), and (25cm sin θ,-25cm cos θ) respectively.

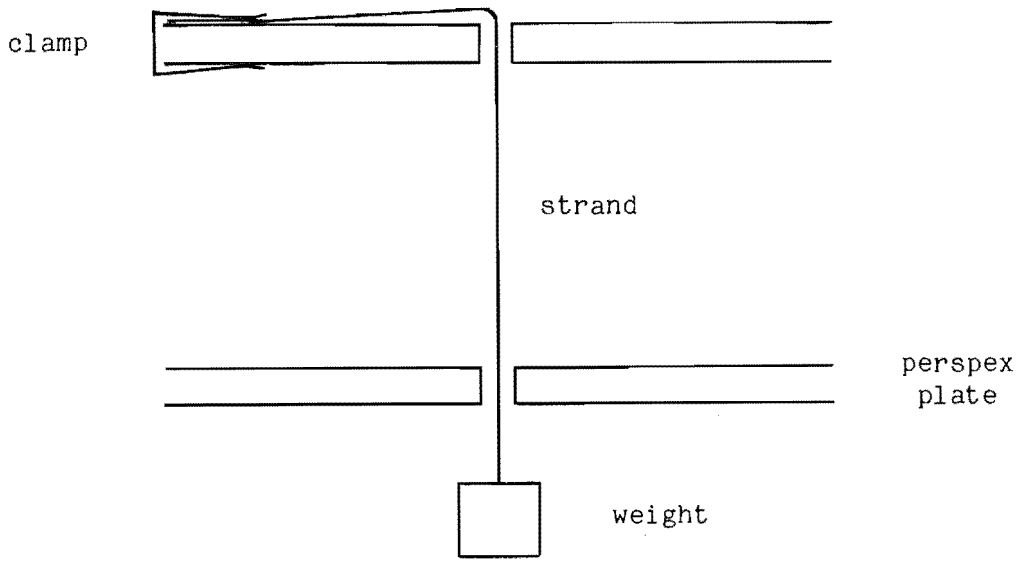
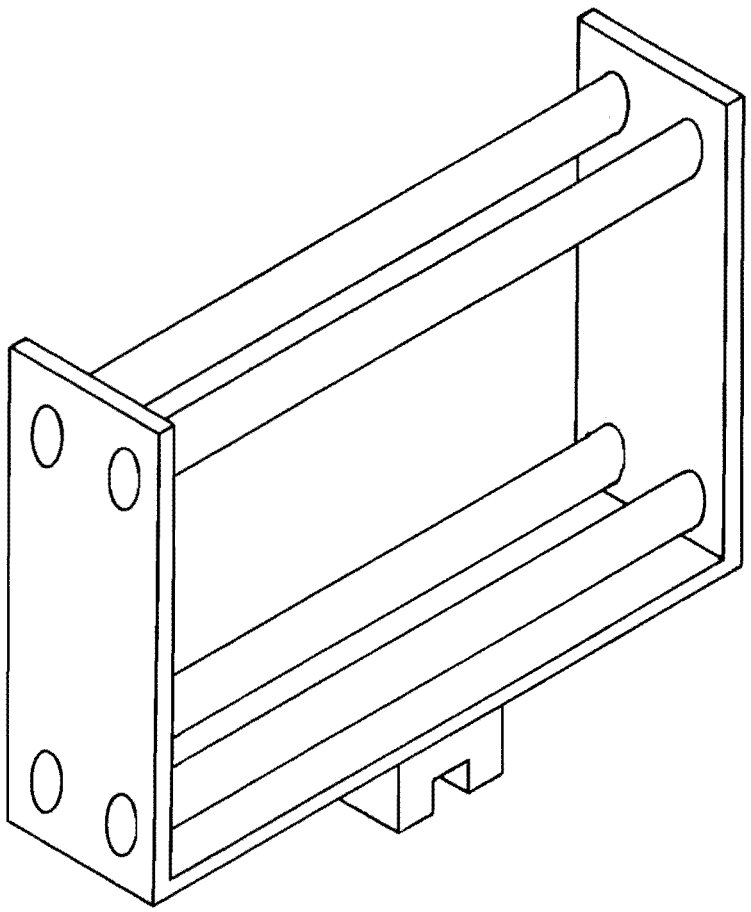


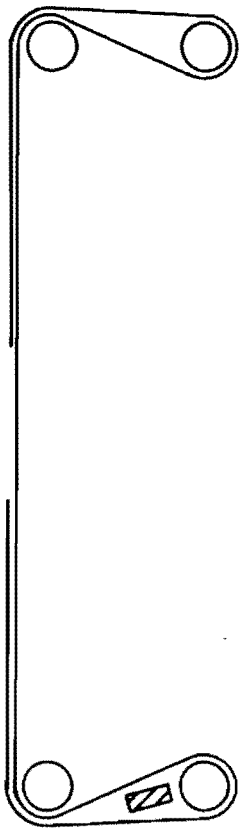
Fig. 5.3 Elevation of plate object holder. After Robinson (1982 p128).

shaped aluminium bracket. The bracket is attached to a base, upon which it may be swivelled. The holder is mounted in the scanner by locating a slot cut in the base on a rail attached to the scanner. A locking screw in the base secures the holder in the selected position. The middle portion of each rod is threaded with a pitch of 0.4 mm. Grid objects are constructed by threading the holder in the manner shown in Fig. 5.4(b) and allowing the strands to settle in the grooves. The strands are fastened by gluing them to one of the rods. The strands may be given additional tension by twisting the perspex strip shown in Fig. 5.4(b) from the horizontal position shown to a vertical one. This method ensures that the strands are parallel and straight. A disadvantage of the grid holder is that because of the intricacy of the threading operation a separate holder is required for each object.

Robinson (1982) considers the effect of wave front distortions introduced by inhomogeneities in the reference medium. In the scanner, the inhomogeneities are constructed by filling buckets with animal tissue. The buckets, which are called 'tissue holders', have windows of 0.12 mm thick



(a)



(b)

Fig. 5.4 Grid object holder.
(a) Isometric view.
(b) Cross-section illustrating the threading method. The shaded rectangle is a cross-section of the perspex strip referred to in the text.

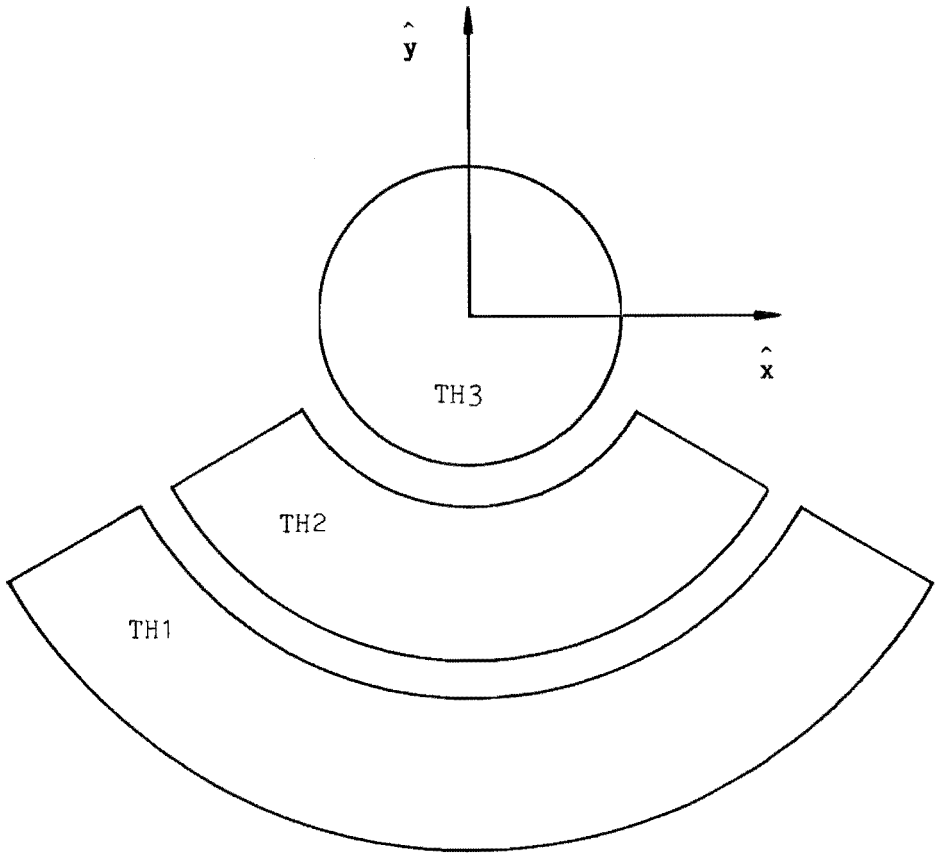


Fig. 5.5 Horizontal cross-section of the tissue holders.

	Inner radius	Outer radius
TH1	20 cm	24 cm
TH2	4 cm	8 cm
TH3		4 cm

mylar in them to allow the ultrasound to pass through the bucket. The buckets are positioned between the object region and the scattered path transducer. The critical dimensions of these buckets are given in Fig. 5.5. Robinson (1982) uses the tissue holders TH1 and TH2. In some imaging situations the inhomogeneities of the reference medium completely surround the object. This is particularly true of medical applications. The tissue holder TH3 was designed and built to enable these situations to be modelled. The main difficulty in designing such a holder is that both a quantity of tissue and an object have to be supported in close proximity. The finally chosen design is shown in Fig. 5.6.

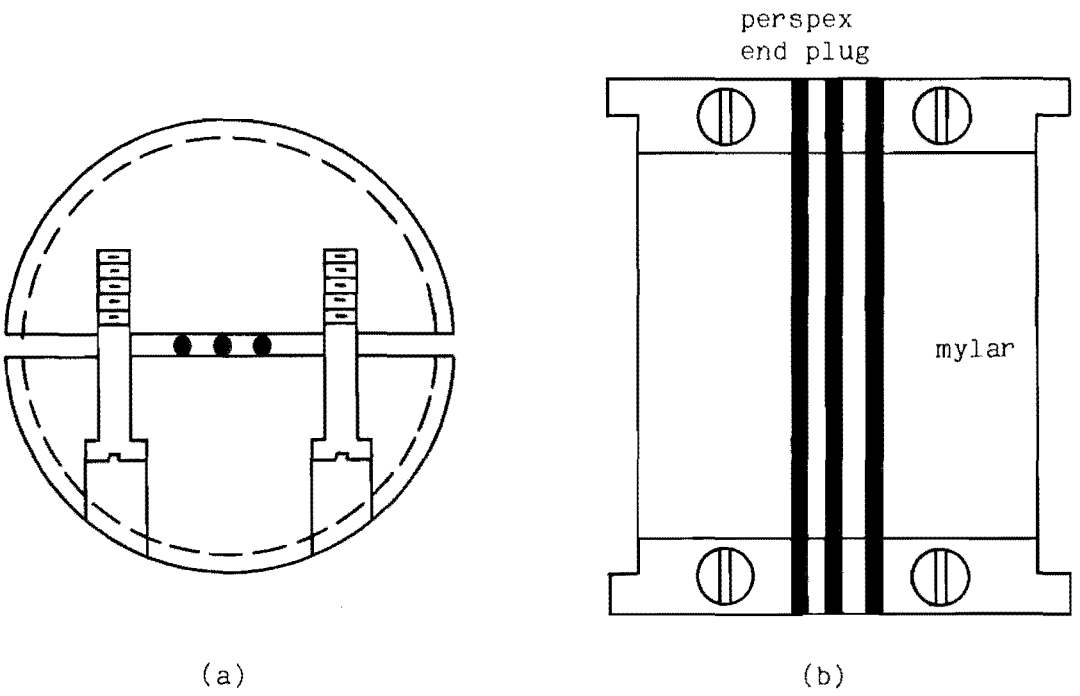


Fig. 5.6 The tissue holder TH3.
(a) Horizontal cross-section of an end plug. The dashed circle shows the location of the mylar wall. The dots are a cross-section of the strands clamped between the two halves of the end plug.
(b) Elevation of the holder.

The tissue is contained in a cylinder. The wall of the cylinder is formed from a 0.12 mm thick mylar strip, the ends of which are overlapped 3 mm and glued with ADOS F3 contact adhesive. The join, which could scatter ultrasound if carelessly fabricated, is oriented at $\theta = 180°$ to place it on the edge of the incident beam. Two end plugs made from 13 mm thick perspex form the ends of the cylinder. Perspex is used so that the interior of the holder may be examined for the presence of air bubbles. Each plug has a tapered recess to locate the wall conveniently and accurately. Also, each plug is constructed in two semicircles to allow objects to be supported within the holder. Two screws allow the semicircles to be clamped together. Strands may be threaded between the halves and clamped in place. This method is employed when TH3 is used in conjunction with the grid

holder. The holder TH3 is simply clamped onto the grid, so that the grid supports the tissue holder. This method allows linear arrays surrounded by a distorting medium to be scanned.

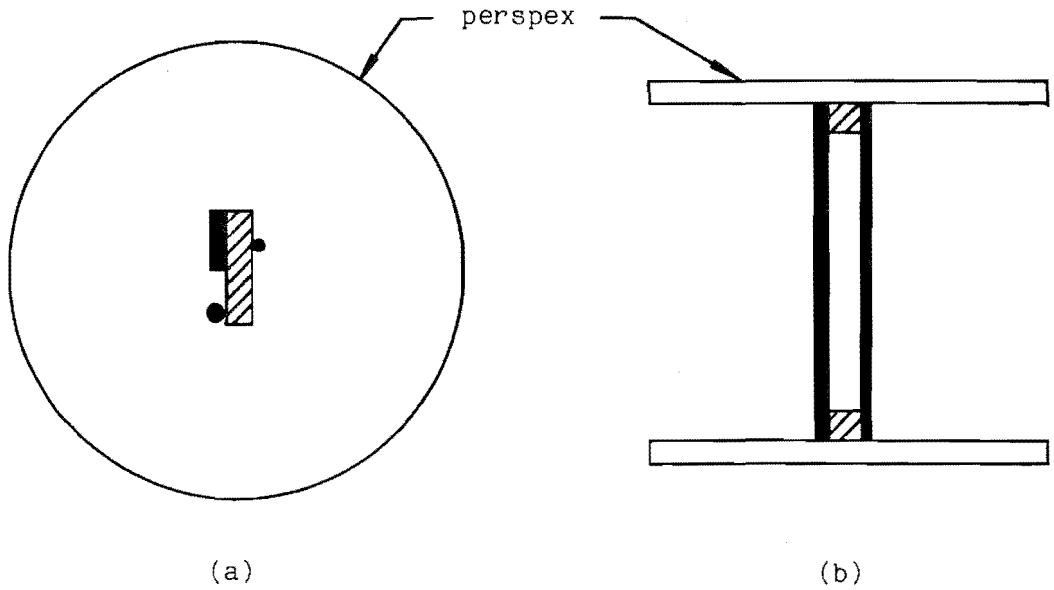


Fig. 5.7 An object supported by the cassette method.
The cassette is placed within TH3.
The shaded regions represent a cross-section of a brass block. Solid filled areas correspond to a cross-section of the object.
(a) Horizontal cross-section.
(b) Elevation.

A 'cassette' method has been devised to allow a wider range of objects to be placed within the tissue holder TH3. The method is illustrated in Fig. 5.7. The object consists of a strip and two wires. Each end of the three components of the object is soldered to a brass block which is in turn glued by Araldite to a 3 mm thick perspex disc. This arrangement, referred to as the 'object cassette', may be placed within the holder TH3. The cylindrical symmetry of the cassette allows it to be placed at any orientation.

5.2.4 Position Control

The two mechanical scanning motions r, θ (Fig. 5.2) of the scanner are under numerical control (Dunlop 1978 ¶4.7). Each motion is derived from an Evershed Model FDD4 stepper motor. For the angle control, a worm and gear wheel are used to produce a 0.072° angular increment per step. A rack and bevel gear are used to generate a radial scanning motion with an increment of 0.12 mm per step. A block diagram of the control system for one of these motions is shown in Fig. 5.8. The computer in control of the scanner supplies the 'required position' as a 12-bit control word. The scanner maintains a record of its 'current position' in an up/down counter. The required and current positions are compared and the current position is decremented or incremented so as to bring the two one count closer. Concurrently a step in reverse or forward command is sent to the stepper motor. When the required and current position counts are the same the clock is disabled.

Mechanical play in the position control system is compensated for by setting the required position to one less count than that actually required when moving in the reverse direction. This is called the 'backlash correction'.

Before using the system, the current position counter must be made to correspond to the actual position of the scanner motion within its range of travel. At each end of the range there is a limit switch which controls the current position counter. The switch at the low end of the range clears the counter, that at the high end sets it. The correspondence of the actual position to the value of the current position counter is established by assigning the zero count to the low end of the range.

A 'position setup procedure' is used to establish the correspondence when the scanner is switched on in an unknown position. The computer sets the current position counter to 4095 and then outputs a required position of 8. The scanner responds to this by starting a sequence of 4087 steps in reverse. However, since the actual number of steps in the range of travel for the angle and radius is 3974 and 2809 counts respectively the lower limit switch must be encountered before the sequence of 4087 steps is completed. At that moment, the current position counter is cleared and the correspondence between actual and current positions is established. The scanner then counts 8 steps forward to the required position.

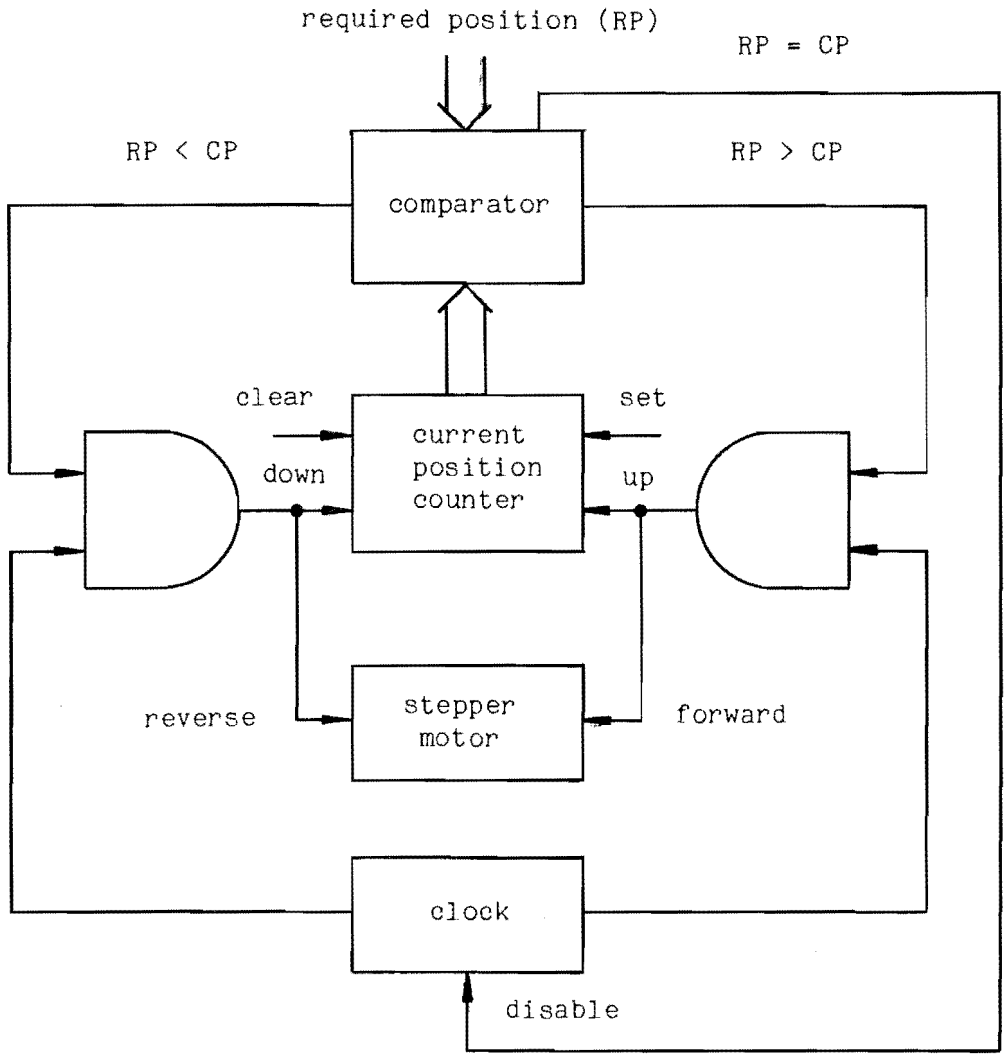


Fig. 5.8 Numerical control system.

5.2.5 Signal Processing

When the scanner has been set to the required position, the transducers are used to make a scattering measurement. Perhaps the simplest form of measurement would involve the transmission of a continuous wave (CW) signal. A disadvantage of such a procedure is that the reverberations set up in the tank would contribute to the measurement and thus complicate its interpretation. A better method is to transmit a linear frequency modulated (FM) sweep and to employ sufficient signal processing in hardware and software to be able to infer the result for quasi-monochromatic insonification.

A block diagram of the scanner hardware used for signal processing is shown in Fig. 5.9. The timing reference is provided by a quartz crystal oscillator which produces a clock signal of frequency f_{SP},

$$f_{SP} = 39062 \text{ Hz} . \tag{5.9}$$

In response to a request from the computer, a linear FM sweep $a_T(t)$ is generated, i.e.

$$a_T(t) = \begin{cases} \cos[\omega_0 t] , & t \leq 0 \\ \cos[(\omega_0 + \tfrac{1}{2}\dot{\omega}t)t] , & t > 0 \end{cases} \tag{5.10}$$

where ω_0 and $\dot{\omega}$ are the sweep parameters, with values

$$\omega_0 = 1.06 \times 10^7 \text{ rad/s} \tag{5.11}$$

and

$$\dot{\omega} = 4.04 \times 10^8 \text{ rad/s}^2 . \tag{5.12}$$

In fact (5.10) does not represent the transmitted waveform for 0 ms < t < 5 ms because of the limitations of the circuitry employed. The signals from transducers R and S are put through gain control stages, which are under computer control, and then multiplied by the transmitted waveform. Assuming a path delay of τ_S, the received signal from the scattered path transducer $a_S(t|\tau_S)$ is

$$a_S(t|\tau_S) = A_S(\tau_S)\, a_T(t - \tau_S) \tag{5.13}$$

where $A_S(\tau_S)$ is a path dependent amplitude factor. The corresponding multiplier output $a_{TS}(t|\tau_S)$ for the scattered path is given by

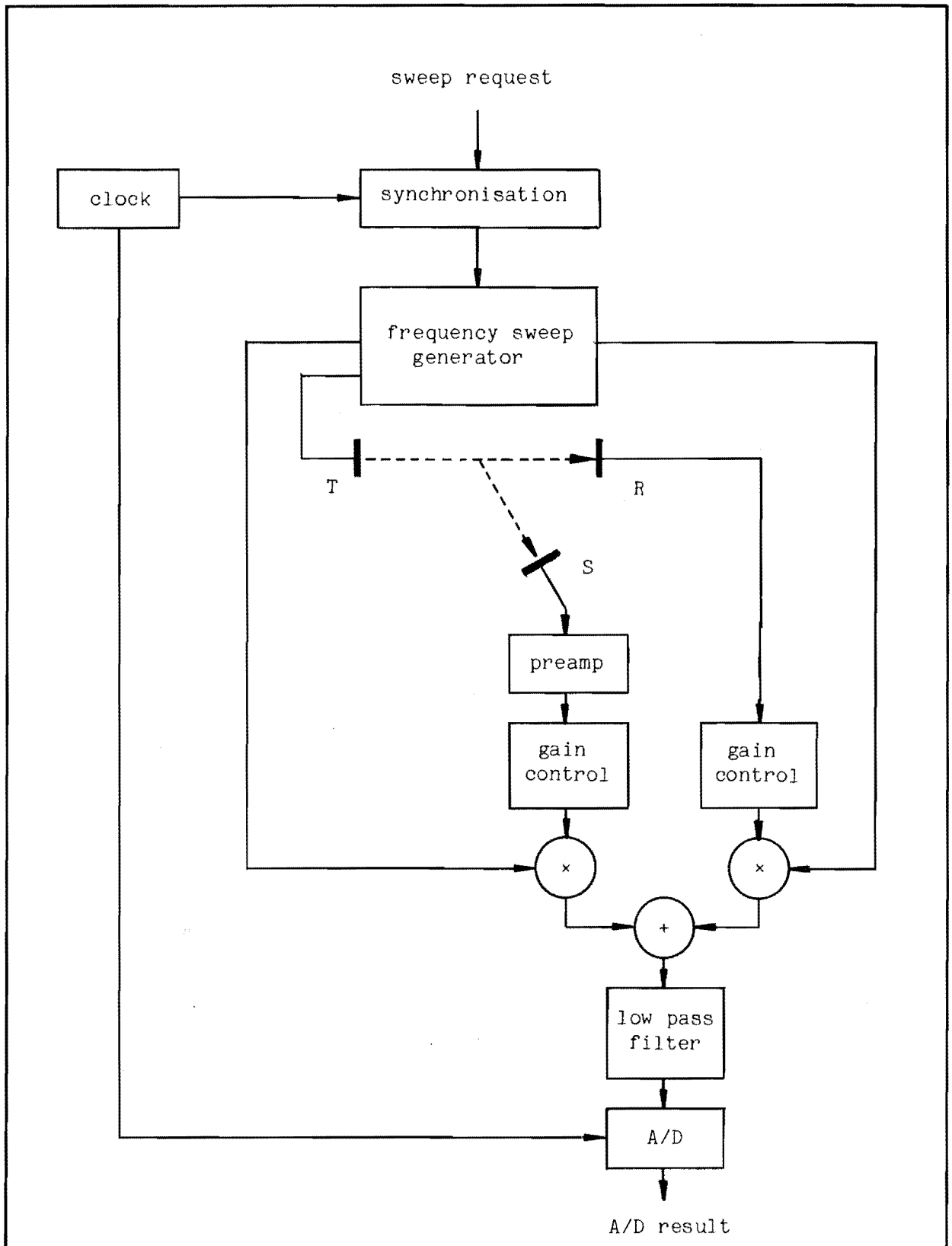


Fig. 5.9 Signal processing hardware. Dashed and solid arrows correspond to ultrasonic and electrical signal paths respectively.

$$a_{TS}(t|\tau_S) = a_T(t)\, a_S(t|\tau_S)\, G_S \tag{5.14}$$

$$= A_S(\tau_S)\, G_S \begin{cases} \cos[\omega_0 t]\, \cos[\omega_0(t - \tau_S)] & t \leq 0 \\ \cos[(\omega_0 + \tfrac{1}{2}\dot{\omega}t)t]\, \cos[\omega_0(t - \tau_S)] & 0 < t \leq \tau_S \\ \cos[(\omega_0 + \tfrac{1}{2}\dot{\omega}t)t] & \\ \quad \times \cos[[\omega_0 + \tfrac{1}{2}\dot{\omega}(t - \tau_S))(t - \tau_S)] & t > \tau_S \end{cases} \tag{5.15}$$

where G_S represents the effect of the scattered path gain control. The product waveform from the scattered and reference paths are added together to form

$$a_{TSR}(t|\tau_S,\tau_R) = a_T(t)[a_S(t|\tau_S)\, G_S + a_R(t|\tau_R)\, G_R] \,. \tag{5.16}$$

This signal is then low pass filtered to obtain the components corresponding to the difference between the transmitted and received frequencies. For the case when the reference path gain is zero, the result is approximately

$$a_{TSRL}(t|\tau_S,\tau_R) = A_S(\tau_S)\, G_S\, G_L \begin{cases} \cos(\omega_0\tau_S) & t \leq 0 \\ \cos(\tfrac{1}{2}\dot{\omega}t^2 + \omega_0\tau_S) & 0 < t \leq \tau_S \\ \cos(\omega_0\tau_S + \dot{\omega}t\tau_S - \tfrac{1}{2}\dot{\omega}\tau_S^2) & t > \tau_S \end{cases} \tag{5.17}$$

where G_L is the gain of the low pass filter. The accuracy of the approximation is least near the breakpoints $t = 0$ and $t = \tau_S$. The general case for $G_R \neq 0$ may be obtained by superposition. The instantaneous frequency of this signal is

$$d(\phi_{TSRL})/dt = \begin{cases} 0 & t \leq 0 \\ \dot{\omega}t & 0 < t \leq \tau_S \\ \dot{\omega}\tau_S & t > \tau_S \end{cases} \tag{5.18}$$

which is graphed in Fig. 5.10. Temporal spectral analysis of a_{TSRL} therefore allows the path dependent amplitude factor $A_S(\tau_S)$ to be determined. To avoid sweep initiation effects present in (5.18) the signal for $t > \tau_{max}$ should be used where τ_{max} is the maximum time delay to be considered. This constraint is reasonable since it is the time taken for a narrow pulse to propagate along the path. Using the dimensions in the caption of Fig. 5.2 and (5.1) for the velocity of sound in water,

$$\tau_{max} = 253\ \mu s \,. \tag{5.19}$$

The output from the low pass filter is sampled by an A/D converter at the frequency f_{SP} and the 12-bit signed twos complement result is available to the computer. Sampling jitter is avoided by synchronisation logic which ensures that the sweep is always triggered on the trailing edge of the clock signal. Since t = 0 is defined by the start of the sweep, the samples are always taken at the same times. The samples corresponding to the first 5 ms are discarded because the sweep may be nonlinear during this period. When this is done all of the remaining samples correspond to path delays exceeding τ_{max} so the sweep initiation effects described by (5.18) do not need to be considered.

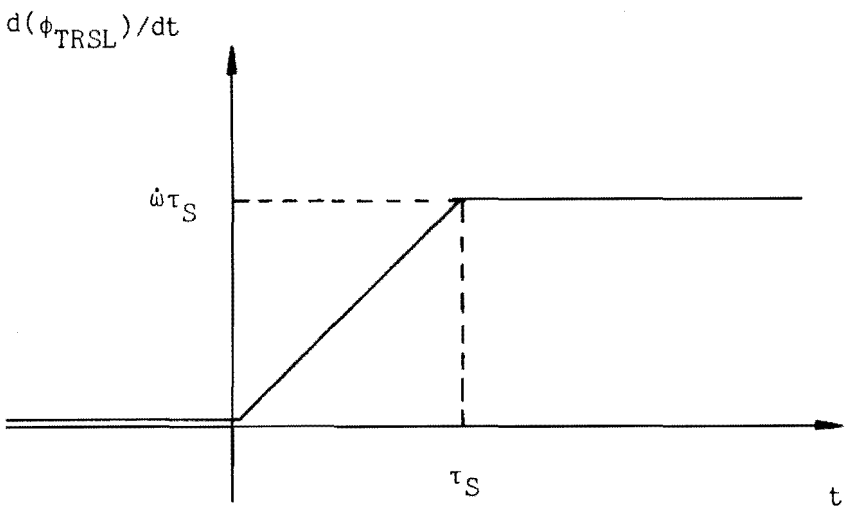


Fig. 5.10 Linear FM sweep signal processing.
The instantaneous frequency of the processed signal is proportional to the scattered path time delay when the reference path gain is zero. When both paths contribute to the result, two frequency components are present in the processed signal.

5.3 INTERFACING CONSIDERATIONS

The design of the scanner-VAX interface is described in this section. The design philosophy employed is presented in ¶5.3.1. Since the scanner had already been interfaced to an EAI-590 computer many aspects of the interface were essentially fixed and these are outlined in ¶5.3.2. A review of the pertinent aspects of VAX input/output (i/o) is given in ¶5.3.3, along with interfacing considerations. Finally, ¶5.3.4 details the assignment of scanner signals to VAX i/o devices.

5.3.1 Philosophy

The goal of the interfacing project was to interface the scanner to the VAX as quickly as possible while minimising the risk of unforseen difficulties extending the time required. Since there is little existing hardware documentation (see Dunlop 1978 Ch. 4; Robinson 1982 Ch. 5) for the scanner, it was felt the latter consideration meant that extensive hardware modifications to the scanner were best avoided. In contrast the i/o devices and associated software for the VAX are well documented. Hence, it was decided to use as many VAX i/o devices as necessary to accomodate the existing signal requirements of the scanner.

Another goal was to design the hardware component of the interface in such a manner as to eliminate the need to write special purpose VAX software to control the i/o devices. A program referred to as a 'driver' is available as part of the VAX/VMS operating system to control each i/o device (Digital 1983). In some circumstances, the driver is unable to control the i/o device in the desired manner and the user is then required to construct a special purpose driver. This is a difficult and therefore costly procedure and the interface was designed to avoid this.

In modern electronic instrumentation, microprocessors (μP) are commonly used for control purposes. A μP could have been used to control the scanner. However, since algorithm development is more readily carried out on a VAX than on any available μP-based system the data obtained would be transferred to the VAX for processing anyway. Thus in this approach two interfaces would have to be designed and built, one from the scanner to the μP and one from the μP to the VAX. On complexity grounds, the μP approach was ruled out.

5.3.2 EAI-590 Implementation

The EAI-590 (EAI) to which the scanner was originally interfaced was a hybrid computer consisting of a digital computer linked to an EAI-693 interface. Both components of the hybrid computer were used to control the scanner. Digital i/o on the EAI-640 was used for sweep and position control (¶5.2.4) and for receiving the sampled sweep (¶5.2.5). Analog output devices on the EAI-580 were used to adjust the gain control sections (¶5.2.5). Consequently, several cables were required to connect the scanner to the computer.

Data acquisition was interrupt-driven. When the received signal was digitised (¶5.2.5) the transfer of the 12-bit result from the scanner to the computer was accomplished by interrupting the computer. In response to the interrupt, the computer executed a routine which read the 12 bits and stored them in memory. This was repeated for each sample taken.

The interrupt line served a dual purpose. In addition to indicating that samples were ready, it indicated that the scanner stepper motors were stationary.

Software for the scanner-EAI interface was written entirely in assembly language and so relied heavily on comments to convey its purpose to the user. This is undesirable since the computer executes the machine instructions in the assembly language program and ignores the comments. The user therefore has no guarantee that the comments are in fact an accurate description of the program.

5.3.3 VAX input/output capability

The hardware configuration of the Department's VAX is shown in Fig. 5.11 (Digital 1980). Two buses - the MASSBUS, and the UNIBUS - and the VAX 11/750 central processing unit (CPU) connect together all devices in the system. The MASSBUS and UNIBUS are connected to the CPU by adapters, the MBA and the UBA respectively. All of the i/o devices suitable for the control of laboratory equipment are connected to the UNIBUS. The piece of equipment being controlled by the i/o device is referred to as the 'user device'. Data flow from the user device to the i/o device is referred to as input, and flow in the opposite direction is referred to as output.

Digital i/o is provided by three DR11-Ks (Digital a) and a DR11-W (Digital b). Each of these is capable of transferring a 16-bit word in either direction between a user device and the UNIBUS. Both of these devices support programmed i/o in which the CPU controls the transfer and

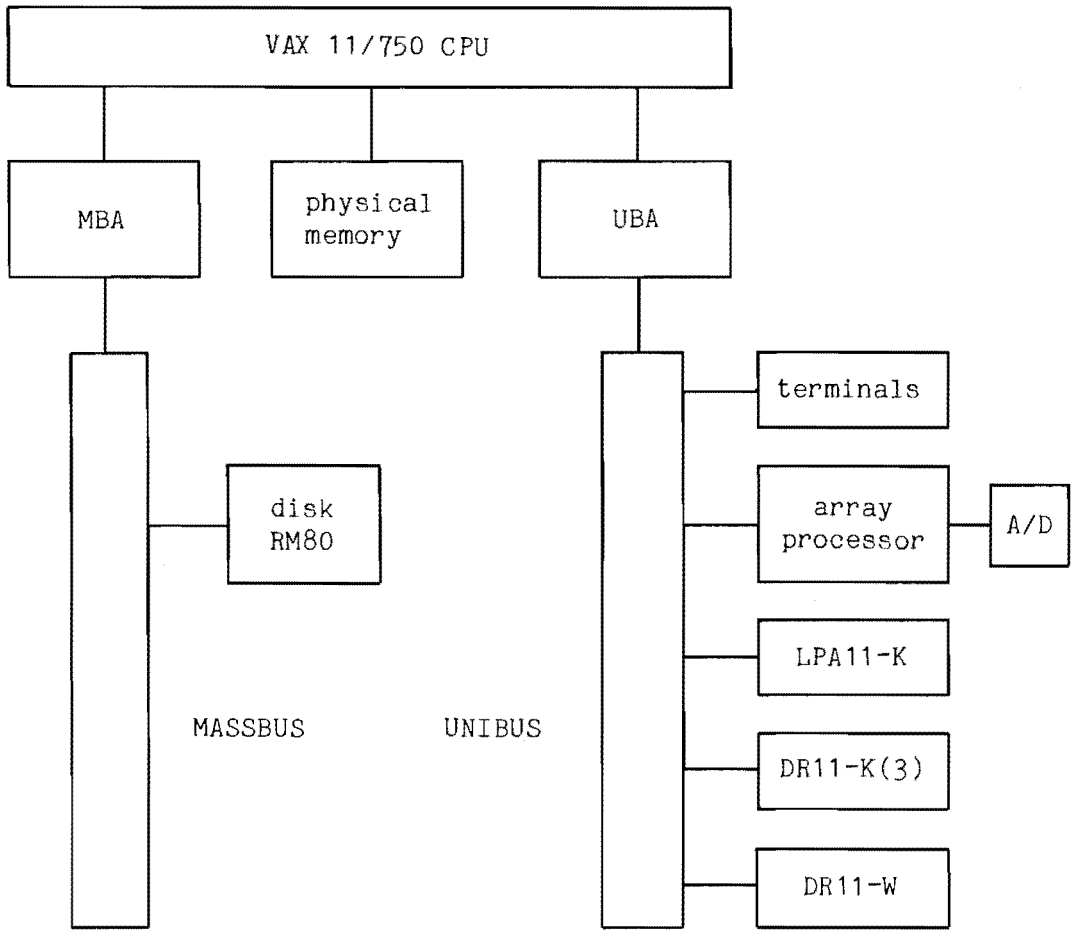


Fig. 5.11 The VAX 11/750 computer of the Electrical and Electronic Engineering Department, March 1984. A number in parentheses indicates the number of devices present.

the word of data is moved between the user device and a CPU register. However, the DR11-W is able to transfer data by direct memory access (DMA). In a DR11-W DMA transfer, data is moved between the user device and memory via the CPU, UBA, and UNIBUS. During the transfer the CPU is able to continue executing instructions.

The 16 input signals for a DR11-W or DR11-K are denoted by DI. Similarly the output signals are denoted by DO. Apart from the 16 input and 16 output signals the DR11-W has several lines which may be used to control or monitor a user device. There are three input lines called STATUS A, STATUS B, and STATUS C, and three output lines FUNCT 1, FUNCT 2, and FUNCT 3. A signal called ATTN may be raised by the user device to cause an interrupt. Data input on the DI lines is accepted when a signal called CYCLE RQA is raised by the user device.

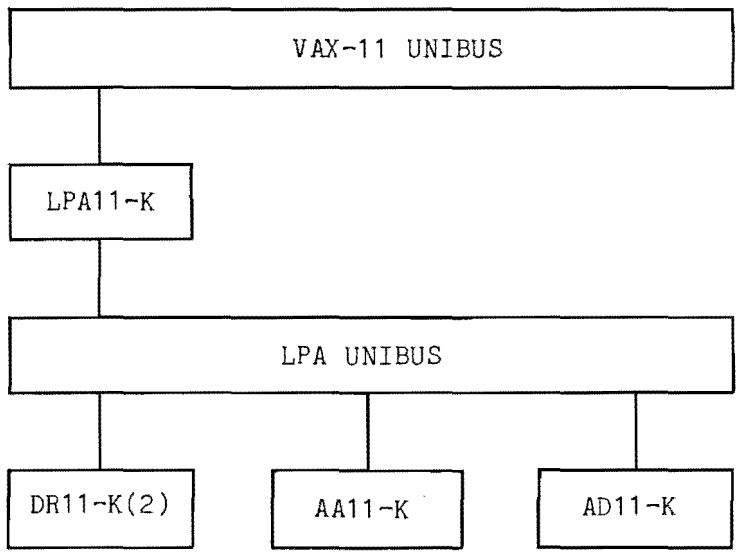


Fig. 5.12 Laboratory peripheral accelerator subsystem.

Additional i/o devices are connected to a separate UNIBUS which is controlled by a laboratory peripheral accelerator (LPA11-K) (Digital c) as shown in Fig. 5.12. The LPA11-K can transfer data to memory by DMA. To the second UNIBUS are connected two more DR11-Ks, an AA11-K (Digital d), and an AD11-K (Digital e). The AA11-K provides 4 channels of digital to analog

(D/A) conversion. Eight channels of analog to digital (A/D) conversion are available on the AD11-K. An additional A/D channel is provided on an array processor which is connected to the main UNIBUS.

The principal difference between the EAI and the VAX as far as interfacing is concerned is that the VAX may be used by several people at once while the EAI was a 'single-user' system. This means that every resource that the VAX has has to be shared. Referring to Fig. 5.11, it is easy to identify the resources concerned. Important examples are storage on disk and in memory, the computing time of the CPU, i/o devices such as terminals and digital i/o devices, and bus bandwidth.

A consequence of the 'multi-user' nature of the system is that an interrupt-driven data acquisition method such as that used on the EAI is impracticable. This is because the CPU needs to respond to each interrupt very quickly in order to capture each sample as it is digitised. On the VAX this can only be accomplished be requiring the machine to idle just prior to the interrupt . This would cause a considerable degradation in response to other users of the machine. A better alternative is to use DMA since the CPU is not required to execute instructions to effect the transfer.

Since the VAX incorporates A/D devices, the question arises as to whether it might be more expeditious to dispense with the A/D section of the scanner and digitise the sweep with one of the A/D devices of the computer, rather than to capture the digitised samples from the scanner with a digital input device. If the A/D attached to the array processor were used for this purpose, spectral processing could be performed on the sweep immediately. One difficulty that arises when an A/D converter within the computer is used is due to the fact that the sweep must be synchronised with the sampling clock. Since the scanner already has the synchronisation logic, it was decided to retain the A/D section of the scanner and use a digital input device.

As mentioned in ¶5.3.1, driver software is available to control most of the i/o devices. Since this software is not designed with a particular user device in mind, the data transfer protocols employed are simple. The dual function of the interrupt line (¶5.3.2) in the scanner is unconventional and in the scanner-VAX interface its function is replaced by three simpler signals.

5.3.4 Signal Assignment

Taking into account the considerations raised in ¶¶5.3.1 through 5.3.3 a set of interface signals was decided upon. At the same time, the signals were assigned to a particular VAX i/o device. The interface signals are listed in Table 5.1. These signals were also used in the scanner-EAI interface with the exception of the MOTOR, FAULT and INTERRUPT signals. These three signals were introduced to simplify the function of the interrupt line.

Table 5.2 specifies the signal assignment. Digital i/o is performed by the DR11-W and a DR11-K. Analog output is provided by two channels on the AA11-K. The DR11-W and the LPA11-K were chosen because they are both capable of DMA and driver software is available for each. The extra signals of the DR11-W are particularly convenient for controlling various scanner functions and for monitoring the status of the scanner. The ATTN signal is supported by the driver software. The driver allows a user supplied routine to be executed in response to an interrupt. The digitised sweep is transferred by DMA to memory at the sampling frequency (5.9) which is well below the maximum transfer rate of approximately 500 000 word/s of a DR11-W in this configuration (see Digital b). The ANGLE and RADIUS signals are transferred in programmed i/o mode, but since they change infrequently this places very little burden upon the CPU. The ANGLE signal is assigned to the DR11-W and the RADIUS to the DR11-K rather than vice versa because when the scanner is being used for far-field measurements the RADIUS signal is not required and therefore the DR11-K need not be used at all. A DR11-K attached to the LPA11-K rather than to the main UNIBUS is used for simplicity. Since the LPA11-K is used to provide analog output, the software for the interface has to control the LPA11-K. It can reasonably be expected that the extra programming required to access a DR11-K attached to the LPA11-K is less than that required to access a DR11-K attached to the main UNIBUS.

5.4 HARDWARE IMPLEMENTATION

To implement the scanner-VAX interface, a certain knowledge of the scanner circuitry was required. However, the documentation of this aspect of the scanner is somewhat sparse (Dunlop 1978 Figs. 4.14, 4.15, 4.17). A wiring list indicating the interconnections between the major components of the scanner was obtained by inspecting the scanner. The results of this exercise are recorded in Appendix A. A considerable effort would be

Table 5.1 The interface signals used by the scanner. The type codes are defined by: DI = digital input; DO = digital output; AI = analog input; AO = analog output. The polarities of the signals which are present on the interface cables are shown in Table 5.2.

Signal Name	Type Code	Description
SETUP	DO	When true the current position counters for both motions are set to 4095.
ANGLE	DO	A 12-bit word specifying the required position for the angle motion.
RADIUS	DO	A 12-bit word specifying the required position for the radius motion.
MOTOR	DI	When true the stepper motors are on, when false the motors are off.
FAULT	DI	When true indicates that a fault condition (5.2.4) is present.
INTERRUPT	DI	A pulse is generated by the scanner on this line when the status of the scanner changes significantly. Significant changes are the motors stopping or a fault occurring.
AGC_SCATTERED	AO	An analog voltage applied to the gain control section for the scattered path to set the gain.
AGC_REFERENCE	AO	An analog voltage applied to the gain control section for the reference path to set the gain.
SWEEP_RQST	DO	When true the computer is requesting the scanner to start a sweep. When false the sweep is reset.
DATA_READY	DI	When this signal goes high, the output from the A/D converter is valid.
A/D	DI	The 12-bit 2s complement result from the scanner's A/D converter.

Table 5.2 Signal assignment. The notation DI(11:0) refers to the bits 11 through 0 of the DI signal. The devices LPA11-K/AA11-K and LPA11-K/DR11-K are connected to the LPA UNIBUS rather than the main UNIBUS. A slash preceding a signal name indicates the ones complement of the signal.

Device	Device Signal	Interface Signal
DR11-W	ATTN	INTERRUPT
	STATUS A	MOTOR
	STATUS B	FAULT
	FUNCT 1	SETUP/
	FUNCT 2	SWEEP_RQST/
	CYCLE RQA	DATA_READY
	DI(11:0)	A/D
	DO(11:0)	ANGLE/
LPA11-K/AA11-K	channel 1	AGC_SCATTERED
LPA11-K/DR11-K	channel 0 DO(11:0)	RADIUS/

required to produce a complete description of the scanner circuitry. However, the information contained in Appendix A is sufficient to allow the more important signals to be found on the scanner. This was sufficient for testing and debugging the scanner-VAX interface.

Very few modifications to the scanner electronics were required to implement the interface. Using the wiring list given in Appendix A, the signals suitable for generating the MOTOR and FAULT signals were found. Buffer circuitry was added to allow these signals to be transmitted to the VAX. Additional circuitry was added to the scanner to generate the INTERRUPT signal. In the EAI implementation the time at which the first sample was taken varied unpredictably. This did not effect the results of Robinson (1982) because the first 195 samples of the digitised sweep are discarded for the reason explained in ¶5.2.5. The clock signal supplied to the A/D conversion circuitry was modified so that the first sample is always taken at the same time.

The gain control sections respond to an analog voltage generated by the AA11-K. This voltage is applied to an opto-isolator and resistor in series within the scanner. The voltage range available from the AA11-K is less than that available on the EAI and so the resistor value was reduced to compensate for this. This alteration necessitated the calibration procedure described in ¶5.6.3.

As remarked upon in ¶5.3.2, the EAI interface required the user to make several connections. To simplify this procedure in the VAX interface all of the cables from the scanner are bound together and anchored to a box at the VAX end of the cables. The cables are connected to sockets provided in the box. All of the connections required for the interface can be made by plugging in the appropriate cables from the VAX into the sockets. The layout of the sockets and their pin assignments are given in Appendix B.

5.5 SUPPORT SOFTWARE

5.5.1 Introduction

A preliminary version of a subroutine package to control the i/o devices which handle the interface signals is presented by Nhunduru (1984). The description contained in this section pertains to an extended and modified package. The way in which the package is combined with the GEM (Chapter 4) to allow interactive control of the scanner is also described.

The structure of the software is explained in ¶5.5.2 and the syntax of the commands used to control the scanner are listed and explained in ¶5.5.3. A data structure used to record the status of the scanner is described in ¶5.5.4. The initialisation procedure for the scanner is given in ¶5.5.5. The algorithms used for position control and sampling are described in ¶¶5.5.6 and 5.5.7 respectively. Finally, in ¶5.5.8 the commands used to collect data when the scanner is used as part of a CAI system are listed.

5.5.2 Software Structure

The user software which controls the interface can be divided into three levels as shown in Fig. 5.13. The user supplies commands to a GEM command module (¶4.8) called the 'scanner>' module. This module responds by calling either 'acquisition' or 'control' routines. Each control routine is responsible for only a small subset of the interface signals.

An acquisition routine carries out a sampling sequence which provides a data set (¶1.5.2) suitable for CAI. The sampling sequence is achieved by calling control routines.

System software is used to manipulate all of the i/o devices used. Driver routines are available for the DR11-W and the LPA11-K. An extra level of sophistication is provided for the LPA11-K by an additional set of 'high level' routines.

5.5.3 Command Syntax

In response to the command 'Help' the scanner> module displays the information shown in Fig. 5.14. This is a concise if somewhat cryptic description of the command syntax used by the module. Strings (¶4.5) commencing with a capital letter are keywords and are entered as is. In this and the following subsections the capital letter of the keyword is retained wherever the keyword is referred to. This helps to prevent confusion between the conventional meaning of the keyword and its meaning as part of a scanner> command. Parameters or qualifiers (¶4.4) for which the user substitutes a particular value are indicated by strings beginning with a lowercase letter. Some of these strings are further defined at the end of the list of commands. These definitions consist of the string being defined, an equals sign, and the definition. In addition, qualifiers are preceded by a slash.

The parameter class (¶4.5) of the parameter determines the set of values allowed for the parameter. Strings representing scalar, vector, or matrix parameters end in '_sca', '_vec', or '_mat' suffixes respectively. If a particular data type (¶4.5) is required (¶4.6), the suffix is preceded by a code to indicate the data type. For example, the codes for integer^2 and real^4 (Tables 4.2 and 4.3) are '_i2' and '_r4' respectively.

The denotation (I) indicates the initial qualifier or value which is assumed by the module when the module is first used. For instance, the initial qualifier for the Interrupt command is /Enable and the initial value for the Radix parameter of the Status command is 'decimal'. The denotation (D) indicates a default qualifier or value which is assumed by the module at the beginning of parsing the command concerned. Optional elements of the syntax are shown in square brackets: e.g. the command [No]Verify has two forms Verify, and NoVerify. Where a choice of one of several options is required this is indicated by listing the options and enclosing the list in braces: e.g. the option 'DR11W' may be chosen for the

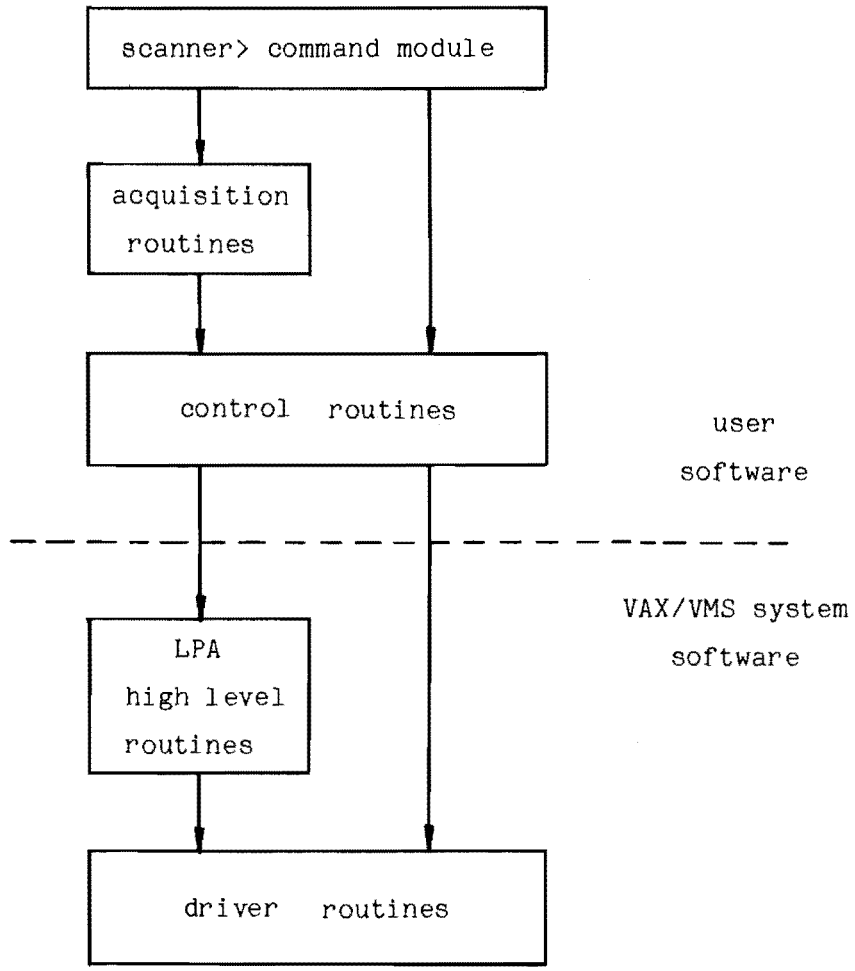


Fig. 5.13 Software structure. An arrow from a block A to a block B indicates that routines in A may call those in B but not vice versa.

```
Ultrasonic scanner commands are:

        AGC     /Retry=retry_i4_sca
                /Scattered/Reference/Sum(I)
                /Minimum=(scat,ref,sum)_min_r4_sca
                /Maximum=(scat,ref,sum)_max_r4_sca
        Allocate/module
        Angle/coordinates               angle_sca
        BSR                             header  acquire file
        Deallocate/module
        Delay                           delay_sca
        Fault/Retry                     retry_i4_sca
        Fault/Noretry(D)                just/print
        Gain/Scattered=gain_scat_i4_sca/Reference=gain_ref_i4_sca
        Initialise/module
        Interrupt/Enable(I)/Disable
        Motors                          just/print
        Position/coordinates            angle_sca       radius_sca
        Reset/module
        Radius/coordinates              radius_sca
        Sample/sweep/AGC(I)             gain_(scat,ref)_i4_sca sample_i2_vec
        Sample/sweep/NoAGC              sample_i2_vec
        Settle/Set                      settle_sca
        Settle/Noset(D)
        Setup/High(I)/Low
        Sintheta                        theta_vec       acquire
        Status                          just/print
                /Radix=radix(I decimal)/Phase=phase(I degrees)
        Sweep/sweep
        Timeout                         dr11w_timeout
        Wait/Notimeout(D)
        Wait/Timeout                    wait_timeout
        [No]Verify
        Zero/Angle=angle_i4_sca/Radius=radius_i4_sca
        coordinates     = (     Index           = parameter is i4,
                                Physical(1)     = parameter is r4)
        module          = (All(D),State,DR11W,LPA11K,Position)
        acquire         = gain_scat_i2_vec gain_ref_i2_vec sample_i2_mat
        sweep           = (Reset,Start(I))
```

Fig. 5.14 The response to the scanner> module Help command.

module qualifier. A list enclosed by parentheses indicates that the string in which the list occurs should be repeated with the parenthesised list replaced by each element of the list in turn. As an example of this repetition, the first two parameters of the Sample/AGC command are 'gain_scat_i4_sca gain_ref_i4_sca'. Explanatory remarks, which are not part of the syntax, are preceded by hyphens. The remarks are not terminated in any formal manner. In some cases qualifiers are listed on the lines following the command to which they apply. This is done to make the syntax easier to read or because of lack of space.

5.5.4 The State Table

A data structure, called the 'state table', is used to record the values of various control parameters and the status of the scanner and i/o devices. The Status command displays the contents of the state table. A sample display is shown in Fig. 5.15. The radix of count parameters, such as the required position for the angle motion, may be specified as a command qualifier. A binary radix is convenient during debugging while a decimal radix is more appropriate when the scanner is used for experimental work.

The state table facilitates programming because all of the operating parameters are clearly identified. When the software is being debugged and tested, the status command provides a rapid and complete indication of the status of the interface. A user, unfamiliar with the scanner software, can use the state table as a mental picture or 'user model' of the scanner-VAX interface.

5.5.5 Initialisation

Prior to using the scanner, various initialisation steps must be performed. The commands to accomplish these are Initialise, Reset, and Allocate. The Initialise command performs operations which only need to be carried out once per execution of the image (¶4.3) that the scanner module is part of. The Reset command sets the scanner-VAX system to a known state. Part of this routine performs the position setup procedure which is described in ¶5.2.4. The Reset command is also used if the scanner is switched off and then on again, or after an error has occurred. The Allocate command allocates the i/o devices used by the interface to the executing image. This prevents other users from interfering with these i/o devices. To guarantee that the i/o devices are not accessed by other users, the Allocate command should really precede the Initialise and Reset commands. However, because of a flaw in the design of the LPA11-K system software (as supplied by Digital Equipment Corporation) this is impossible. The flaw is present in the high level routines which in effect employ a second user to reset the LPA11-K. This means that if the LPA11-K is already allocated, it cannot be reset.

5.5.6 Position Control

A scanner position is specified in one of two coordinate systems. In the 'Index' system the angle and radius are each specified by an

```
Ultrasonic scanner status on 20-FEB-1985 20:22:17

    Radix:                  decimal          Initialised:    yes
    Verify:                 no
    Position control:
                            Angle                            Radius
       Index:               0                                100
       Physical:            -176.76 degrees                  12 mm
       Zero:                2455                             0
       Increment:           .072 degrees                     .12 mm
       Backup:              8                                8
    Setup:                  low
    Interrupt event flag:   62               Interrupt:      disabled
    Motors:                 on               Motor count:    0
    Fault:
       Present:             yes              Fault count:    0
       Retry limit:         10               Retry count:    0
    Gains:                                   scattered:      0
                                             reference:      0
    Sample event flag:      63               Sweep:          reset
    Sample delay:           195              Sweep reset:    1 ms
    DR11-W:
       Datapath:            direct           Mode:           user device
       Device:              XAA0:            Initialised:    yes
       Channel:             176              Timeout(s):     10
       ODR:                 64535            Function:       0^B
    LPA11-K:
       Device:              LAA0:            Initialised:    yes
       Interval:
          Desired:          1 ms             Actual:         1 ms
       Clock:
          Rate:             1                Preset:         64536
       Buffer length:       6                Event flag:     61
       Dwell:               1                Delay:          1
    Automatic Gain Control:
                            Minimum                          Maximum
       Scattered:           35%                              45%
       Reference:           35%                              45%
       Sum:                 50%                              90%
       Retry limit:         2                Retry count:    0
       Retry failures:      0
    Settling time:          10 s             Wait timeout:   40 s
    Control-C:
       Wait:                no               Acquisition:    no
```

Fig. 5.15 The response to the scanner> module Status command.

integer. The control routines (¶5.5.2) output the integer as the required position (¶5.2.4). The second system, called the 'Physical' system, uses real numbers to specify the angle and the radius in radians and metres respectively. Each of these numbers is converted to the Index system by dividing by the appropriate step increment and adding on the Index coordinates of the origin in the Physical system. The origin is specified by the Zero command.

Three commands are available to set the scanner to a particular position. The coordinate system which is used for specifying the position is selected by a command qualifier. The Angle and Radius commands output the required position and then return control to the scanner> module so that another command is requested. The scanner may not finish stepping to the required position for several seconds. The Position command outputs the required position and then waits for the stepper motors to stop.

5.5.7 Sampling

A sweep may be triggered and sampled by the Sample command. The samples are stored in the vector specified in the command. A certain number of samples at the beginning of each sweep may need to be discarded because of sweep nonlinearity (¶5.2.5). This number is set by the Delay command. Automatic gain control (AGC) can be selected with a command qualifier to allow the full range of the A/D converter to be utilised.

The AGC is applied to three signals. The 'scattered' signal is obtained by setting the gain of the reference path channel to a minimum. The 'reference' signal is obtained with the scattered path channel set to a minimum. The 'sum' signal is obtained by setting the gain control sections to the gain values found by applying AGC to each of the scattered and reference signals.

The AGC operates by sampling sweeps and adjusting the gain control sections (¶5.2.5) until the reference, scattered, and sum signals meet certain criteria. The criterion is that the maximum absolute value of the samples should lie between a minimum and a maximum value. These values are specified independently for the scattered, reference, and sum signals by the AGC command.

The method by which the gain values are adjusted depends on the signal which is being sampled. For the reference and scattered signals, the gain value is obtained by successive approximation. If this constraint cannot be satisfied, the action taken depends on whether the peak sample

value is too low or too high. If it is too high the AGC operation is considered to have failed. If it is too low a warning message is issued but the AGC operation continues. For the sum signal the gain values are not altered. If the constraint is not satisfied the entire AGC operation is repeated. This is called a 'retry'. The maximum number of retries permitted is specified with the AGC command.

5.5.8 Data Acquisition

When the scanner is used as part of a CAI system based upon right scattering (¶1.8.5), only two commands in addition to the initialisation commands of ¶5.5.5 need to be used. The Sintheta command obtains a data set (¶1.5.2) from the scanner and stores it in memory. The BSR command transfers the data set from memory to a disk file.

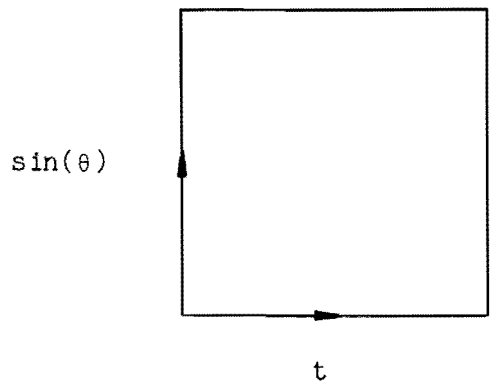


Fig. 5.16 The data set image. The image coordinates are t and sin(θ). Each row of the image is a sampled sweep.

The data set consists of sampled sweeps obtained for several values of the angle θ. The values are chosen to give equal increments in the variable u, which is defined by (1.241) and is called the 'aperture coordinate'. The data set may be considered to be an image according to the definition (1.14). The image coordinates are t and sin θ as shown in Fig. 5.16.

Two integer^2 vectors (¶4.5) and an integer^2 matrix are used to store the data set in memory. One vector is used to store the gain values found by the AGC algorithm for the scattered path, and the other stores the corresponding values for the reference path.

The sampled sweeps are stored in the matrix with one sweep per row. The number N_s of samples in each sweep and the number N_u of angles used are determined by the size of the matrix.

The BSR command creates a disk file in the format used by Robinson (1982). In this format the matrix size is fixed:

$$N_s = 1024 \tag{5.20}$$

and

$$N_u = 64 \ , \tag{5.21}$$

and the aperture is assumed to extend from the angle θ_{min} to θ_{max} where

$$\theta_{min} = -60° \tag{5.22}$$

and

$$\theta_{max} = 60° \ . \tag{5.23}$$

5.6 TESTING AND CALIBRATION

5.6.1 General Procedure

The scanner-VAX interface was debugged and tested in several stages, starting with the scanner> module. Even with the scanner disconnected the control routines can be partially tested. This is possible because these routines update the state table to reflect the actions which they are supposed to have performed. The state table can therefore be inspected with the Status command to check that the control routines are interpreting their arguments correctly. By monitoring the i/o signal pins of the i/o devices, the actions of the control routines can be further verified. The next stage is to connect the scanner to the VAX and to check that the scanner responds to each signal correctly. This is a straightforward procedure for all but the control routine which is used to control the DMA transfer of digitised samples from the scanner to the VAX. The method employed to test this routine is described in ¶5.6.2. The final stage is to test the acquisition routines by acquiring a data set.

A calibration procedure, made necessary by the modified gain control circuitry described in ¶5.4, is outlined in ¶5.6.3.

5.6.2 DMA Testing

Testing a DMA link between two computers is usually straightforward because each computer has software to alter and display portions of its memory. However, in the scanner, samples are transferred by DMA as soon as they are obtained and special equipment is required to display the samples before they are transferred. Fig. 5.17 shows a simple method of accomplishing this. A signal generator, not shown in the figure, drives the sweep request line with a square wave having a 500 μs period. The synchronisation logic (¶5.2.5) then raises the sweep signal in phase with the master clock signal MCLOCK. The ADCLOCK signal used to control the A/D converter is shown in Fig. 5.17(b). A sample is taken on every positive edge of the signal. The sweep is used to trigger the timebase of the oscilloscope. The timebase signal is a positive-going ramp and is connected to the input of the A/D converter. Therefore, the samples obtained by the A/D converter are of the timebase signal and their values can be read off the oscilloscope screen. To execute a DMA transfer of samples to the VAX, the sweep request line is disconnected from the signal generator and connected to the computer. The Sample command is then used to request a sweep and to store the resulting samples. These samples may be compared with the values obtained from the oscilloscope screen to test the DMA transfer.

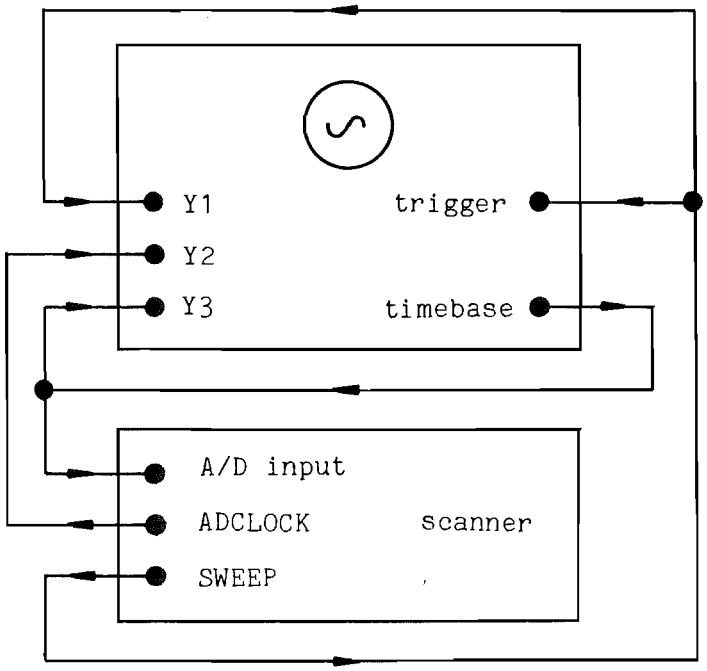
5.6.3 Gain Control

The relationship between the digital control word sent to the AA11-K and the gain applied to the signal in the gain control section is non-linear. This relationship, which is called the 'gain control characteristic', is determined by measurement. The hardware modification, described in ¶5.4, to the gain control section altered the characteristic. The characteristic is required so that the true amplitude may be recovered. Fig. 5.18 shows the modified characteristic for the scattered and reference paths.

5.7 CONCLUSIONS

To test the operation of the scanner-VAX configuration as a whole, measurements were made of the three-strand target used by Robinson (1982

(a)



(b)

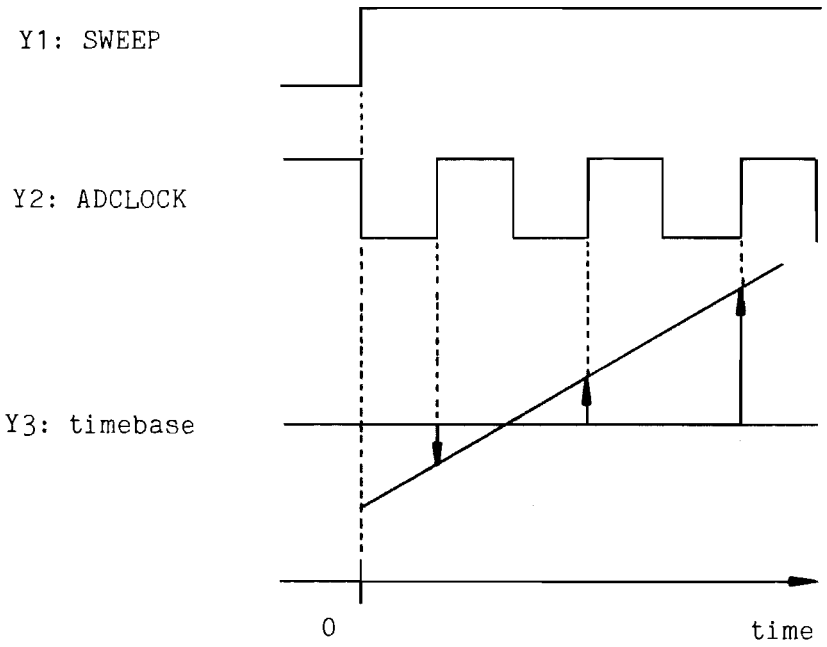


Fig. 5.17 DMA testing.
(a) Equipment setup.
(b) Oscilloscope waveforms.

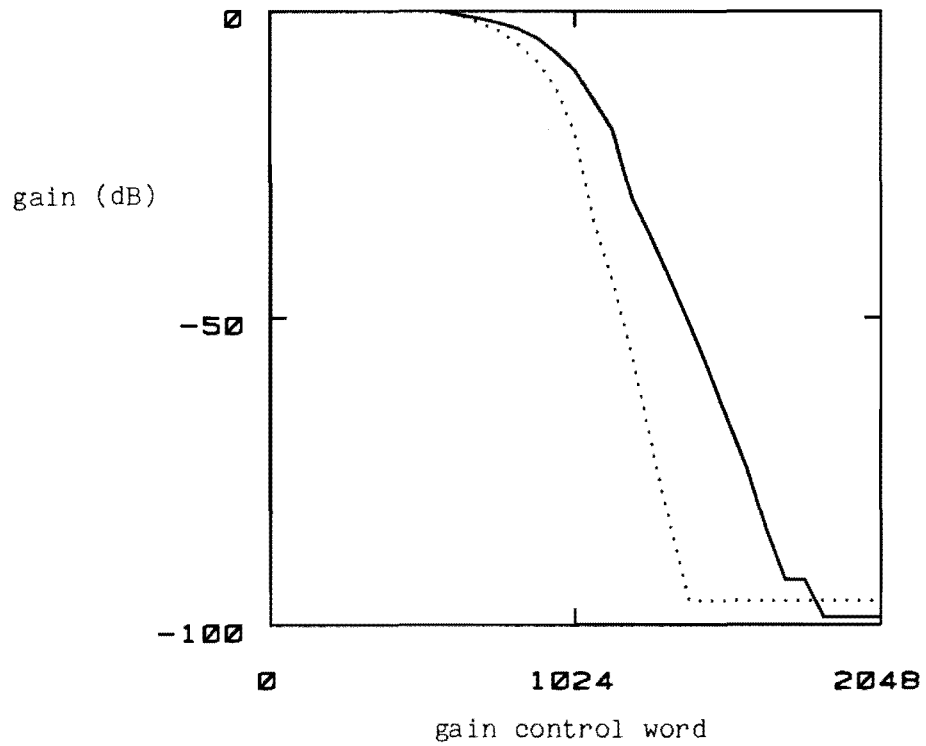


Fig. 5.18 Gain control characteristics measured at 2MHz.
Solid line: scattered path.
Dotted line: reference path.
The gain G is defined by

$$G = 20 \log_{10}(A_{out}/A_{in})$$

where A_{in} is the amplitude of the input to the gain control section and A_{out} is the output amplitude. The plotted gain value is relative to the gain obtained with a gain control word of zero.

Fig. 5.5(b)). The image reconstructed from the measurements (Nhunduru 1984 Fig. 5.13) was comparable to that obtained by Robinson (1982 Fig. 7.14(a)), thus indicating that the scanner-VAX configuration functions correctly as a CAI system.

The scanner-VAX interface is easier to use than the scanner-EAI interface for several reasons. For instance, less time is needed to connect the scanner to the VAX than was required to connect it to the EAI, and there is less likelihood of making a wrong connection. Measurements are easier to make because the scanner can be controlled interactively. Also, complicated data acquisition procedures can be developed rapidly with the aid of GEM command files (¶4.7).

The images of ultrasonic objects presented in Chapters 6 and 7 were reconstructed from measurents made while the scanner was interfaced to the EAI. The results of a comprehensive set of measurements made with the scanner interfaced to the VAX are discussed in ¶10.2.

6. THE SHIFT-AND-ADD PRINCIPLE

6.1 INTRODUCTION

The shift-and-add principle is a simple technique for compensating for speckle distortions (¶1.9). It was conceived in the context of astronomical imaging (Bates and Cady 1980) but has also found application in ultrasonic imaging (Bates and Robinson 1981). In this chapter the basic principle is described and illustrated. New results obtained when the scatterers are extended objects are presented. In ¶6.2 the method used to display two-dimensional images in this thesis is described. A method for generating two-dimensional speckle images is presented in ¶6.3. Examples of shift-and-add images are given ¶6.4. The relationship of shift-and-add to an earlier variant called 'LWH' processing (Lynds et al. 1976) is discussed in ¶6.5 and examples of the latter are presented. Shift-and-add reconstructs objects which contain an isolated bright point and some ways such a point can arise in biological objects are discussed in ¶6.6. For objects which do not contain a single isolated bright point, the reconstructed image is distorted. The nature of the distortion and a simple model to predict it are described in ¶6.7. An interpretation of shift-and-add as an approximate inverse method is given in ¶6.8. The results reported in this chapter are summarised in ¶6.9.

6.2 IMAGE DISPLAY

The two-dimensional images presented in this thesis are processed before being displayed to compensate for the characteristics of the display technique. The display method represents ten levels of intensity for each pixel in the image. The image is normalised before display. Next, the image is thresholded, i.e. all of the pixels with a magnitude of less than d_ε are set to zero. This helps to remove low-level noise and makes detail in the image more apparent. Finally, the image magnitude is quantised to the ten levels available. The quantisation maps image magnitudes between d_0 and d_1 into ten ranges of equal size.

The display method consists of drawing a 'spiral' of a length corresponding to the quantised magnitude. The 'spiral' has a finite line width and so the area of the 'spiral' is proportional to the quantised image magnitude.

The parameters d_0 and d_1 may be chosen to enhance the display. For instance, when $d_0 = 0$ and $d_1 = 1$ image detail of magnitude 0.09 would normally be quantised into the lowest range and be displayed as white. By reducing the value of d_0 the image detail can be quantised into the second to lowest range, and so rendered visible.

The first image coordinate (usually x) is displayed across the page and the second (usually y) is displayed up the page. For the images obtained from experimental data, the x and y axes are oriented with respect to the incident wave and the measurement aperture as shown in Fig. 5.2. The x and y coordinates for these images range from -7 mm to 7 mm.

6.3 SPECKLE SIMULATION PROCEDURE

6.3.1 Method

A method for generating an ensemble $\{s_m\}$ of speckle images of an object f has been devised and programmed by the author. A random and uncorrelated phase function $\phi_m^{(uc)}$ is generated using a pseudo-random number generator. The phases are uniformly distributed in the interval $[-\phi_{max}, \phi_{max})$ where ϕ_{max} is arbitrary. The phase may be constrained to be conjugate symmetric, i.e.

$$\phi_m^{(uc)}(\mathbf{r}) = -\phi_m^{(uc)}(-\mathbf{r}) \quad . \tag{6.1}$$

An 'uncorrelated' speckle PSF $h_m^{(uc)}$ is calculated from the uncorrelated phases $\phi_m^{(uc)}$ according to

$$H_m^{(uc)} = \exp\left(j\phi_m^{(uc)}\right) \tag{6.2}$$

and

$$h_m^{(uc)} = \mathcal{F}^{-1}\left(H_m^{(uc)}\right) \quad . \tag{6.3}$$

The next step is to multiply the uncorrelated speckle PSF by the 'seeing disc' σ to form a 'correlated' speckle PSF $h_m^{(c)}$, i.e.

$$h_m^{(c)} = \sigma h_m^{(uc)} \quad . \tag{6.4}$$

The reason that $h_m^{(c)}$ is referred to as correlated is that the multiplication by σ is equivalent to convolving $H_m^{(uc)}$ with $\mathcal{F}(\sigma)$ (1.63). Filtering a random variable such as $H_m^{(uc)}$ causes it to become correlated (Bracewell 1978 p.328). The correlation is over a region in Fourier space, the size

of which is determined by the extent of $\mathcal{F}(o)$. The region is called a 'seeing cell'. The multiplication by the seeing disc concentrates the image energy near the origin. This is as it should be, since if the phase of $H_m^{(c)}$ was perfectly correlated, i.e. constant, then the correlated speckle PSF would be a δ function (¶1.3.6). Speckle images of the object are calculated from

$$s_m = \mathcal{F}^{-1}\left(F\, H_m^{(c)} \right) . \qquad (6.5)$$

Note that (6.1) to (6.5) simulate coherent speckle formation, i.e. s_m is a complex quantity with an amplitude and a phase. No contamination is added explicitly to the speckle images. However, depending on how the speckle images are processed by a reconstruction algorithm, they may or may not be effectively contaminated. This curious state of affairs arises because the convolution $f \odot h_m$ is performed by an FFT (¶1.8.2). The convolution is therefore calculated cyclically (Bracewell 1978 p.362) which produces a periodic result. If the reconstruction algorithm uses this knowledge and extends the speckle image using its periodicity, then there is no effective contamination. However, if the speckle image is extended with zeros (¶1.8.1) and aliasing (Bracewell 1978 p.377) is present, then the speckle image is effectively contaminated.

The simulation procedure was verified by generating speckle images of simple objects using different seeing discs. A seeing disc consisting of a δ function was used to simulate perfect seeing. Since the Fourier transform of a δ function is a constant, the phases of $H_m^{(c)}$ are constant for this case. A uniform seeing disc was used to simulate very poor seeing. In this case, the phases $\phi_m^{(c)}$ are uncorrelated. A Gaussian seeing disc given by

$$o(\mathbf{r}) = \exp\left[-j\pi(|\mathbf{r}|/2)^2\right] \qquad (6.6)$$

was used to simulate seeing conditions in between these two extremes. This seeing disc was used for all of the simulation results presented in this thesis.

An alternative method for calculating a correlated speckle PSF is to filter the phase $\phi_m^{(uc)}$ directly, rather than to filter $H_m^{(uc)}$. Robinson (1982 ¶4.6) uses a Markov chain to achieve this. Such a procedure does not allow the shape of the autocorrelation function of the sequence of phases, and hence of the seeing disc, to be readily specified. Also, the method is restricted to one dimension. Wesolowski (1983) uses a more general procedure which allows the autocorrelation to be specified.

Fourier transforms are not employed but computational efficiency is enhanced by employing a two stage finite impulse response filter instead of a single filter. An advantage of the author's simulation procedure over that of Robinson (1982) is that two-dimensional speckle images may be simulated. Another useful aspect of the author's simulation method is that in general the data do not need to be windowed (Harris 1978) before the Fourier transform is calculated by the FFT (¶1.8.2). This is because when correlated phases are employed the image amplitude at the edge of the image is very small.

6.3.2 Histograms

The first order statistics of the speckle images of several objects were investigated by calculating histograms. Four objects consisting of a δ, a Gaussian, a rectangle, and a triangle function were considered. The objects and typical speckle images are shown in Fig. 6.1. The speckle images of a δ and a rectangle function are similar since in both cases the amplitude varies rapidly with position. The speckle image for the triangle function and the Gaussian are smooth slowly varying functions and are not in fact 'speckly'. The reason for the difference lies in the different natures of the objects. Both the δ and rectangle objects are discontinuous functions while the Gaussian and the triangle functions are continuous. When the speckle PSF is convolved with a continuous function, the rapid fluctuations of the PSF are smoothed out. However, with a discontinuous function, the rapid fluctuations are preserved.

The image amplitude is a nonstationary stochastic process, because the statistics of the amplitude vary with position. Hence, the histogram p of a function g of the amplitude A depends on the position x and is therefore a function $p(x,g(A))$ of two variables. It can thus be displayed as a two-dimensional image. Histograms for various functions of the speckle image amplitude consisting of the magnitude, phase, real part, and imaginary part for a δ function object are shown in Fig. 6.2. Only one-dimensional results are shown here since the statistics of the two-dimensional case are found to be similar. The phase is uniformly randomly distributed. The magnitude is more likely to have a higher value near the centre of the image. Note that this is to be expected in the case, for example, when a Gaussian seeing disc is employed because the image energy tends to be concentrated in the centre of the image. The real and imaginary parts have the same statistics, which is consistent with the uniform random phase. The phases are not constrained to be conjugate

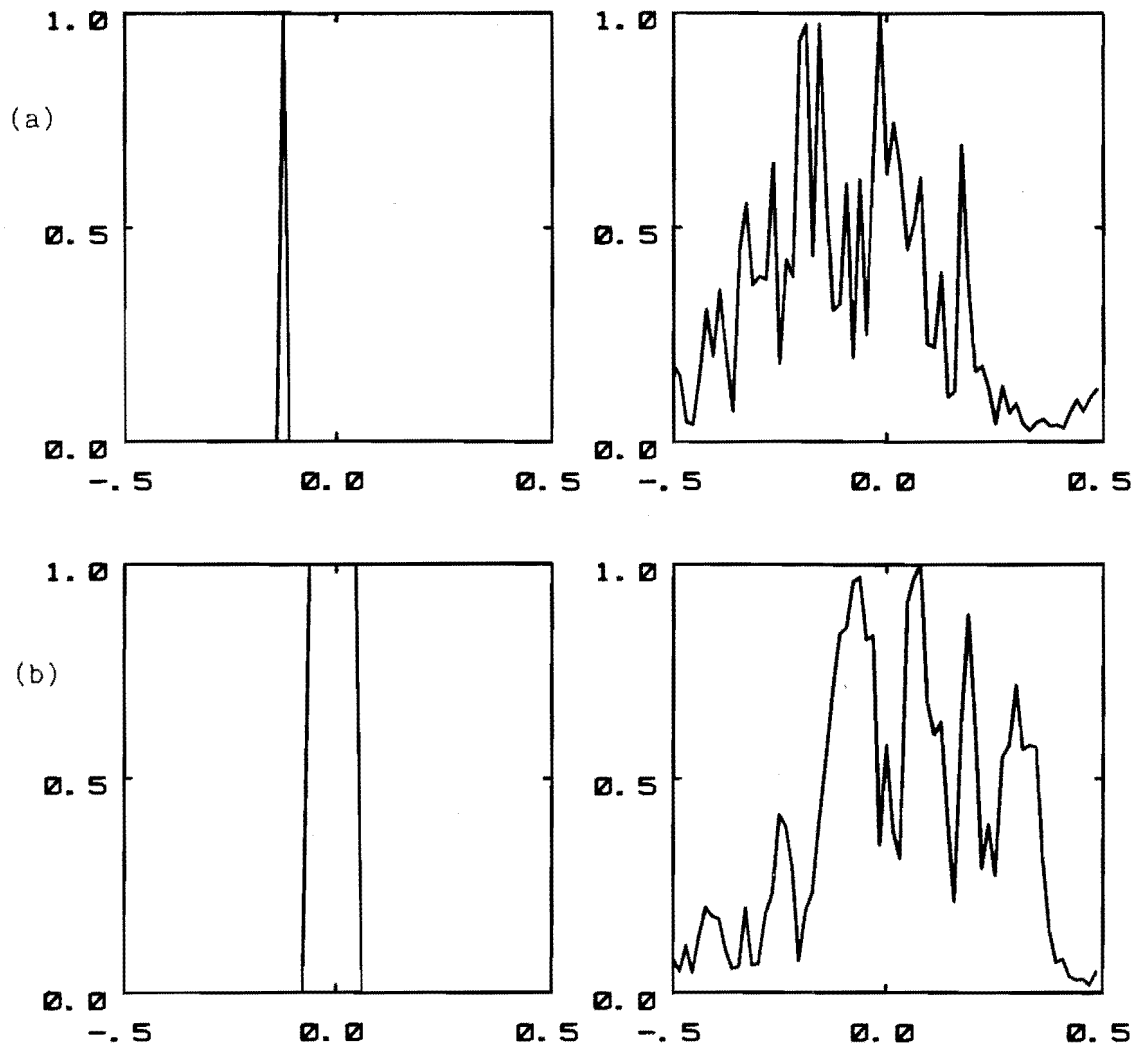


Fig. 6.1 Speckle images of simple objects.
Image magnitudes are displayed.
Left column: object f.
Right column: speckle image s_1.
(a) δ. (b) Rectangle. (c) Triangle.
(d) Solid line: gaussian. Dotted line: seeing disc.

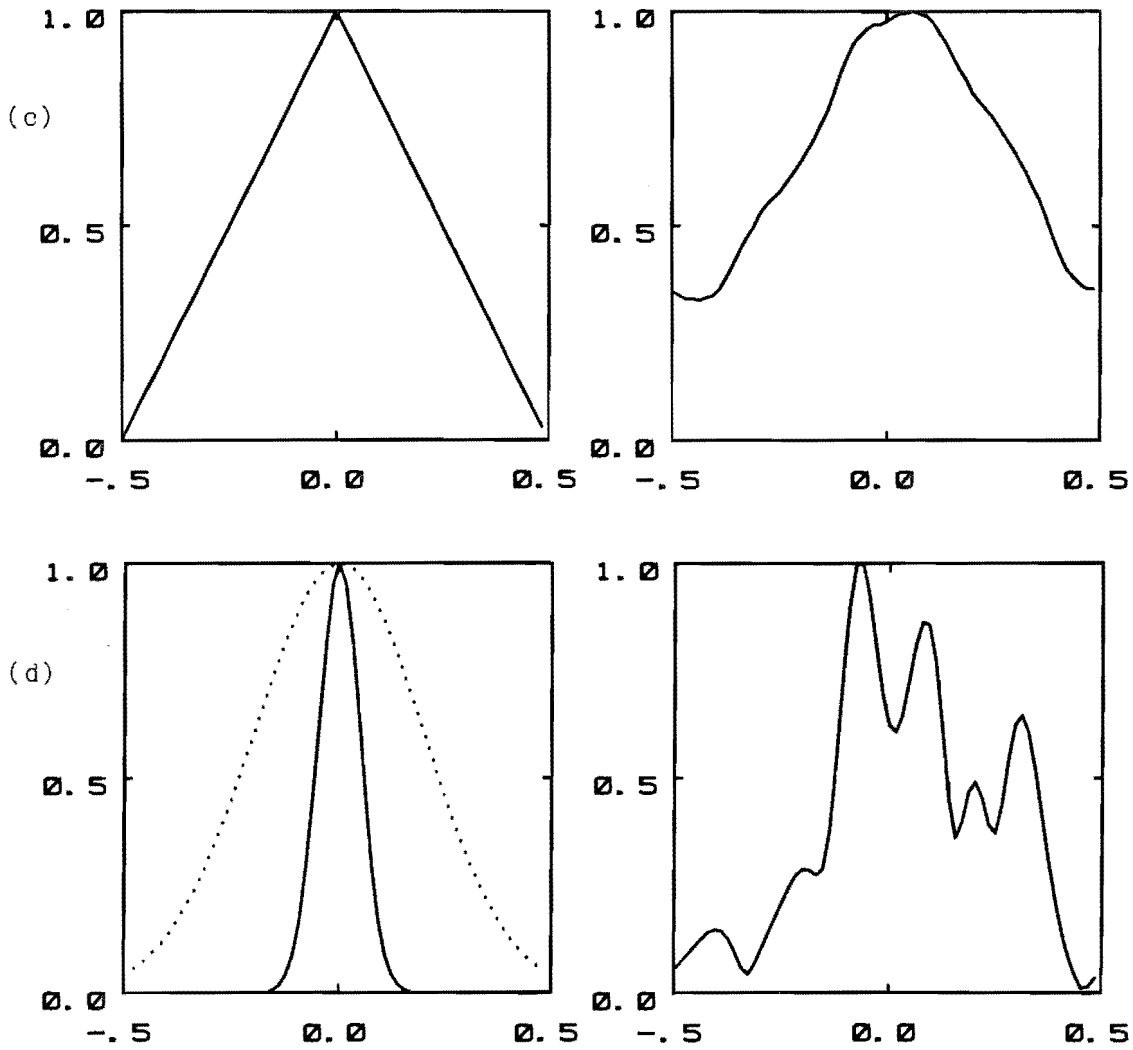


Fig. 6.1 (cont.)

symmetric, hence the real and imaginary parts have similar statistics rather than the imaginary part being identically zero.

Histograms of the magnitude of speckle images of a rectangle, triangle, and Gaussian function objects are shown in Fig. 6.3. The overall shape of the histograms is similar. The only exception is for the triangle function and this is due to its width being effectively somewhat wider than the seeing disc. For all of the other objects (which are narrower than the seeing disc) the statistics are very similar. This effect is invoked in the simple probability model of a speckle image presented in ¶6.7.

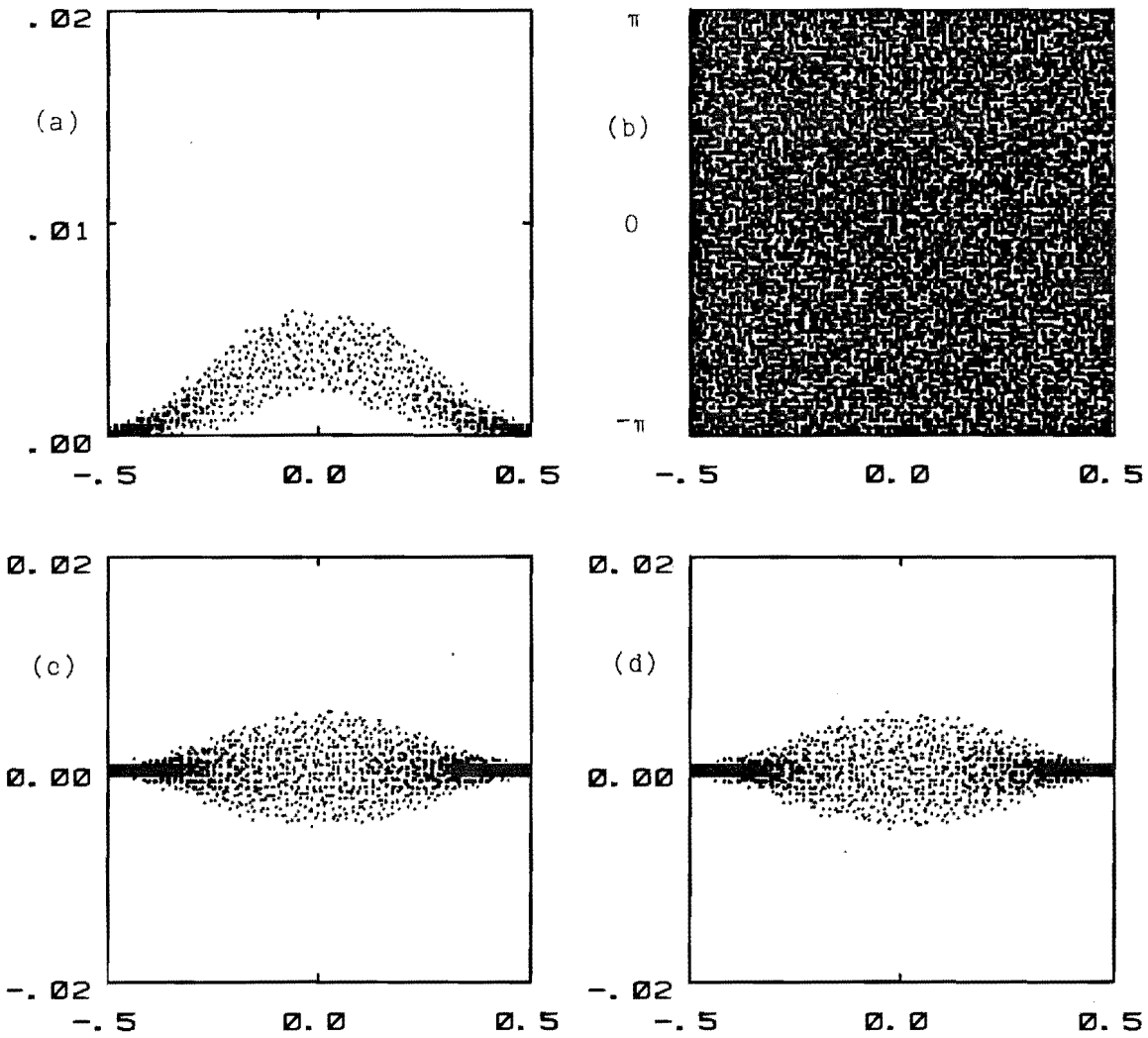


Fig. 6.2 Histograms of speckle images of a δ function object. The speckle image amplitude is A. See ¶6.2 for a description of the display algorithm.

Display parameters: $d_0 = 0$; $d_\varepsilon = 0$; $d_1 = 0.5$.

(a) p(x,|A|) (b) p(x,phase(A))

(b) p(x,re(A)) (c) p(x,im(A))

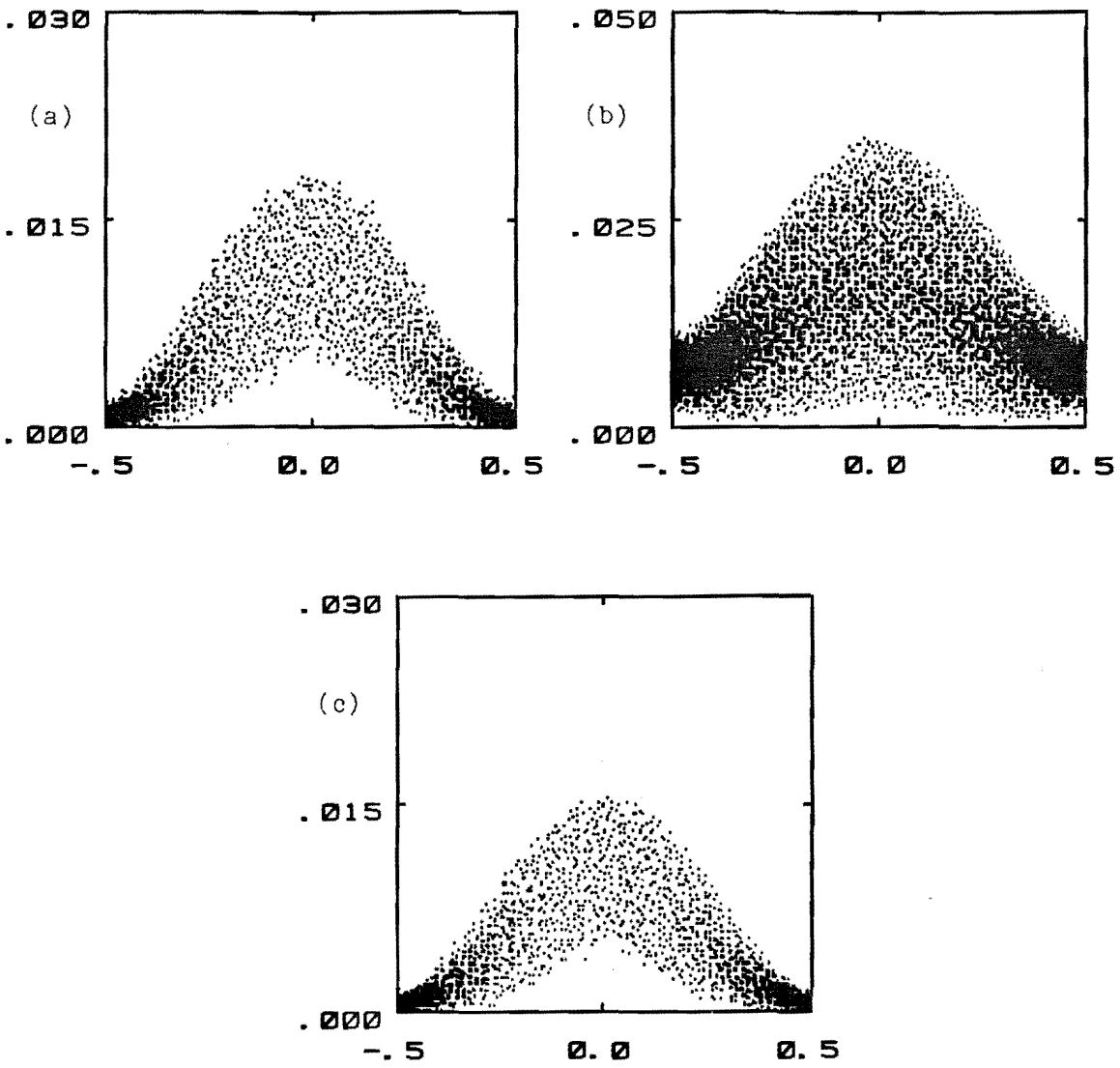


Fig. 6.3 Histograms $p(x, |A|)$ of the magnitude of speckle images of some simple objects.
Display parameters: $d_0 = 0$; $d_\varepsilon = 0$; $d_1 = 0.5$.
(a) Rectangle. (b) Triangle. (c) Gaussian.

6.4 SHIFT-AND-ADD AVERAGES

The shift-and-add principle was conceived as a solution to the speckle imaging problem in the context of optical astronomy (Bates and Cady 1980). The principle has also been applied to ultrasonic imaging by Bates and Robinson (1981) and Minard et al. (1985).

6.4.1 A Simple Model for the Speckle PSF

Consider the function h defined by

$$h(\mathbf{r}) = \sum_l h_l \, \delta(\mathbf{r} - \mathbf{h}_l) \tag{6.7}$$

where $\{h_l\}$ is a set of complex amplitudes and $\{\mathbf{h}_l\}$ is a set of position vectors. Note that the subscript is the letter l and not the digit 1. All of the points specified by $\{\mathbf{h}_l\}$ are within a distance d of the origin, i.e.

$$|\mathbf{h}_l| < d \, . \tag{6.8}$$

An equivalent expression for h which employs the operator T defined by (1.31) is

$$h = \sum_l h_l T_{\mathbf{h}_l}(\delta) \, . \tag{6.9}$$

In this subsection it is shown that h is a valid representation of a general speckle PSF when the h_l and $\mathbf{h}_l$ are random.

The Fourier transform H of h is found, after invoking (1.59), to be

$$H(\mathbf{u}) = \sum_l h_l \exp(j2\pi \mathbf{u}\cdot\mathbf{h}_l) \, . \tag{6.10}$$

Fourier space may be partitioned into regions $\{T_m\}$ for which the argument of the complex exponential in (6.10) does not vary significantly. Let $\mathbf{u}_m$ denote a particular point within the region T_m and assume that the linear dimensions of T_m are of order D. For a point $\mathbf{u}$ in T_m write

$$\mathbf{u} = \mathbf{u}_m + \boldsymbol{\varepsilon} \, . \tag{6.11}$$

Then

$$\begin{aligned} \exp[j2\pi(\mathbf{u}\cdot\mathbf{h}_l)] &\approx \exp(j2\pi \mathbf{u}_m\cdot\mathbf{h}_l)\,(1 + j2\pi\boldsymbol{\varepsilon}\cdot\mathbf{h}_l) \\ &\approx \exp(j2\pi \mathbf{u}_m\cdot\mathbf{h}_l) \end{aligned} \tag{6.12}$$

if

$$|2\pi\boldsymbol{\varepsilon}\cdot\mathbf{h}_l| \ll 1 \, . \tag{6.13}$$

Now since $|\mathbf{h}_l| < d$, (6.13) is satisfied if

$$|\varepsilon| \ll 1/2\pi d \, . \tag{6.14}$$

Hence D may be chosen to be

$$D = 1/d \, . \tag{6.15}$$

Define a real amplitude A_m and a phase θ_m for the region T_m by

$$A_m \exp(j2\pi\theta_m) = H(\mathbf{u}_m) \ . \tag{6.16}$$

From (6.10), A_m and θ_m can be seen to be the result of a random walk (Goodman 1975) and hence are random variables. In conclusion then, the formula (6.9) for a speckle PSF is consistent with $H(\mathbf{u})$ having a random amplitude and phase, which are correlated over a distance of 1/d.

6.4.2 Speckle Image Structure

A speckle image of the object f is formed according to

$$s_m = h_m \odot f + c_m \tag{6.17}$$

where h_m is a speckle PSF and c_m is a contamination term. When the model (6.9) for the speckle PSF is substituted into (6.17), the speckle image is seen to be a sum of randomly located replicas of the object with random amplitudes:

$$s_m = \sum_l h_{m,l} T_{\mathbf{h}_{l,m}}(f) + c_m \ . \tag{6.18}$$

6.4.3 Ideal Shift-and-Add

A possible strategy for solving the 'speckle imaging problem', i.e. reconstructing f from $\{s_m\}$, is to average the speckle images in such a way that replicas of the object reinforce each other. This is the essence of shift-and-add. One way to do this is to take the brightest replicas from each speckle image and superpose them. Let the value of l for which $|h_{m,l}|$ is largest for a given m be denoted by L(m), i.e.

$$|h_{m,L(m)}| = \max_l \{|h_{m,l}|\} \ . \tag{6.19}$$

The replica corresponding to $\mathbf{h}_{m,L(m)}$ is the brightest one. The superposition of the brightest replicas is called the 'ideal' shift-and-add image f_{id}. It is defined by

$$f = \left\langle \frac{T_{\mathbf{h}_{m,L(m)}}(s_m)}{h_{m,L(m)}} \right\rangle_m \tag{6.20}$$

where the angular brackets in (6.20) denote an average over the ensemble of speckle images, e.g.

$$\langle s_m \rangle = \frac{1}{M} \sum_{m=1}^{M} s_m \ . \tag{6.21}$$

Upon substituting (6.9) into (6.20), f_{id} may be related to the object through

$$f_{id} = f + c_{id} \tag{6.22}$$

where the 'ideal' contamination c_{id} is given by

$$c_{id} = \Big\langle \sum_{l \neq L(m)} \frac{h_{m,l}}{h_{m,L(m)}} T_{\mathbf{h}_{l,m} - \mathbf{h}_{l,L(m)}}(s_m) + \frac{T_{-\mathbf{h}_{m,L(m)}}(c_m)}{h_{m,L(m)}} \Big\rangle_m . \tag{6.23}$$

The ideal contamination consists of randomly shifted and scaled versions of the fainter replicas and contaminations in each speckle image.

6.4.4 Weighted Shift-and-Add

In practice, f_{id} cannot be used to solve the speckle imaging problem because the $h_{m,l}$ and $\mathbf{h}_{m,l}$ are unknown. However, approximate methods for estimating them can be devised. First, though, some preliminary definitions are required. Denote a location, called a 'brightest' point, for which $|s_m|$ attains its maximum value, by $\mathbf{b}(s_m)$,

$$|s_m[\mathbf{b}(s_m)]| = \max_{\mathbf{r}}\{|s_m(r)|\} . \tag{6.24}$$

The software used to generate the results presented in this thesis determines $\mathbf{b}(s_m)$ by examining an I×J sampled representation $s_{i,j}$ of s_m (¶1.2.5) in the order $s_{1,1}$, $s_{2,1}$, ..., $s_{I,1}$, $s_{1,2}$, $s_{2,2}$, ..., $s_{I,2}$, $s_{1,J}$, $s_{2,J}$, ..., $s_{I,J}$. If s_m has more than one brightest point then the first encountered is used. The amplitude of the brightest point determined by (6.24) is denoted by

$$b(s_m) = s_m[\mathbf{b}(s_m)] . \tag{6.25}$$

Also a 'weighting' function $w^{(n)}$ is required to weight the speckle images according the (complex) amplitude A of their brightest point. The weighting function, defined by

$$w^{(n)}(A) = \frac{|A|^n}{A} , \tag{6.26}$$

where n is an as yet unspecified integer, has two useful properties. Firstly, the A in the denominator normalises the speckle image with respect to its brightest point, and secondly, the numerator allows a speckle image to be given additional weight if it is brighter than its fellows.

The n-weighted shift-and-add image $f^{(n)}_{wsaa}$ is defined by

$$f^{(n)}_{wsaa} = \langle w_m T_{-\mathbf{r}_m}(s_m) \rangle_m \tag{6.27}$$

where

$$\mathbf{r}_m = \mathbf{b}(s_m) \tag{6.28}$$

and

$$w_m = w^{(n)}\big(b(s_m)\big) \ . \tag{6.29}$$

For general objects there is no reason why $f^{(n)}_{wsaa}$ should approximate f_{id} at all. However, when the object contains an isolated bright point B, called the 'reference', the brightest point of a speckle image often corresponds to B. This is sufficient to allow the relative location of a replica of the object to be determined. This replica is then superposed with its fellows in the other speckle images. This may be expressed mathematically as follows. Denote the amplitude of the reference by a and its location by $\mathbf{a}$. The object is

$$f = aT_{\mathbf{a}}(\delta) + g \tag{6.30}$$

where g is denotes that part of the object other than the reference. The image g is smaller in magnitude than the δ function. The speckle images of f are formed by substituting (6.30) into (6.18),

$$s_m = \sum_l a h_{m,l} T_{\mathbf{h}_{m,l}} + \mathbf{a}^{(\delta)} + \sum_l h_{m,l} T_{\mathbf{h}_{m,l}}(g) + c_m \ . \tag{6.31}$$

Since g is small, the brightest point of s_m is usually determined by the first term on the left hand side of (6.31),

$$\begin{aligned} \mathbf{b}(s_m) &= \mathbf{b}\Big[\sum_l a h_{m,l} T_{\mathbf{h}_{m,l}} + \mathbf{a}^{(\delta)} \Big] \\ &= \mathbf{h}_{m,L(m)} + \mathbf{a} \ . \end{aligned} \tag{6.32}$$

(6.32) is true only for most of the $\{s_m\}$, rather than for all of them, because there is the possibility that the first two terms on the left hand side of (6.31) might interfere with each other in such a way as to cause the true brightest point to be fainter than the second term. This discussed in more detail in ¶6.7. Assuming (6.32) to be true for all of the $\{s_m\}$, and considering only the case n = 0,

$$f^{(0)}_{wsaa} = \left\langle \frac{T_{-\mathbf{h}_{m,L(m)}} - \mathbf{a}^{(s_m)}}{a h_{m,L(m)}} \right\rangle_m$$

$$= \left\langle \frac{T_{-a}}{a} \left[\frac{T_{-\mathbf{h}_{m,L(m)}}{}^{(s_m)}}{h_{m,L(m)}} \right] \right\rangle_m$$

$$= \frac{T_{-a}(f_{id})}{a} \qquad (6.33)$$

where (6.20) is invoked. From (6.33), it can be seen that 0-weighted shift-and-add reconstructs the ideal shift-and-add image for objects of the form given by (6.30) but that the overall amplitude and location are unknown if a and **a** are unknown. For the rest of this thesis the reference is assumed to be of unit strength and located at the origin, i.e.

$$a = 1 \qquad (6.34)$$

$$\mathbf{a} = \mathbf{0} \qquad (6.35)$$

unless specified otherwise.

The weighting constant n, determining the form of the weighting function, is selected to suit the imaging context. For instance, the adjusted shift-and-add image f_{asa} defined by Bates and Fright (1982) corresponds to the choice n = 2. For optical speckle images, which are real and positive, this gives additional weight to the brighter speckle images and results in a more faithful image. Simple shift-and-add f_{saa} defined by Minard et al. (1985) corresponds to n = 1 and is employed by Bates and Robinson (1981). The elementary shift-and-add image f_e (Minard et al. 1985) corresponds to n = 0.

A simulation of f_e is shown in Fig. 6.4. Note that the object has a bright reference, and that the contamination is reduced by increasing the number of speckle images averaged. In Fig. 6.4(d) there is a peak corresponding to each peak present in the object. The peaks in f_e are in correct locations but the peaks other than the one corresponding to the reference are slightly fainter (≈10%) than in the object. Also there is an additional 'ghost' peak at x = 0.2. The ghost peak is present because in some speckle images the brightest point corresponds to the object peak at x = -0.2 rather than to the reference. The speckle image is therefore shifted incorrectly, and the reference peak is shifted to x = 0.2 where it gives rise to the ghost peak.

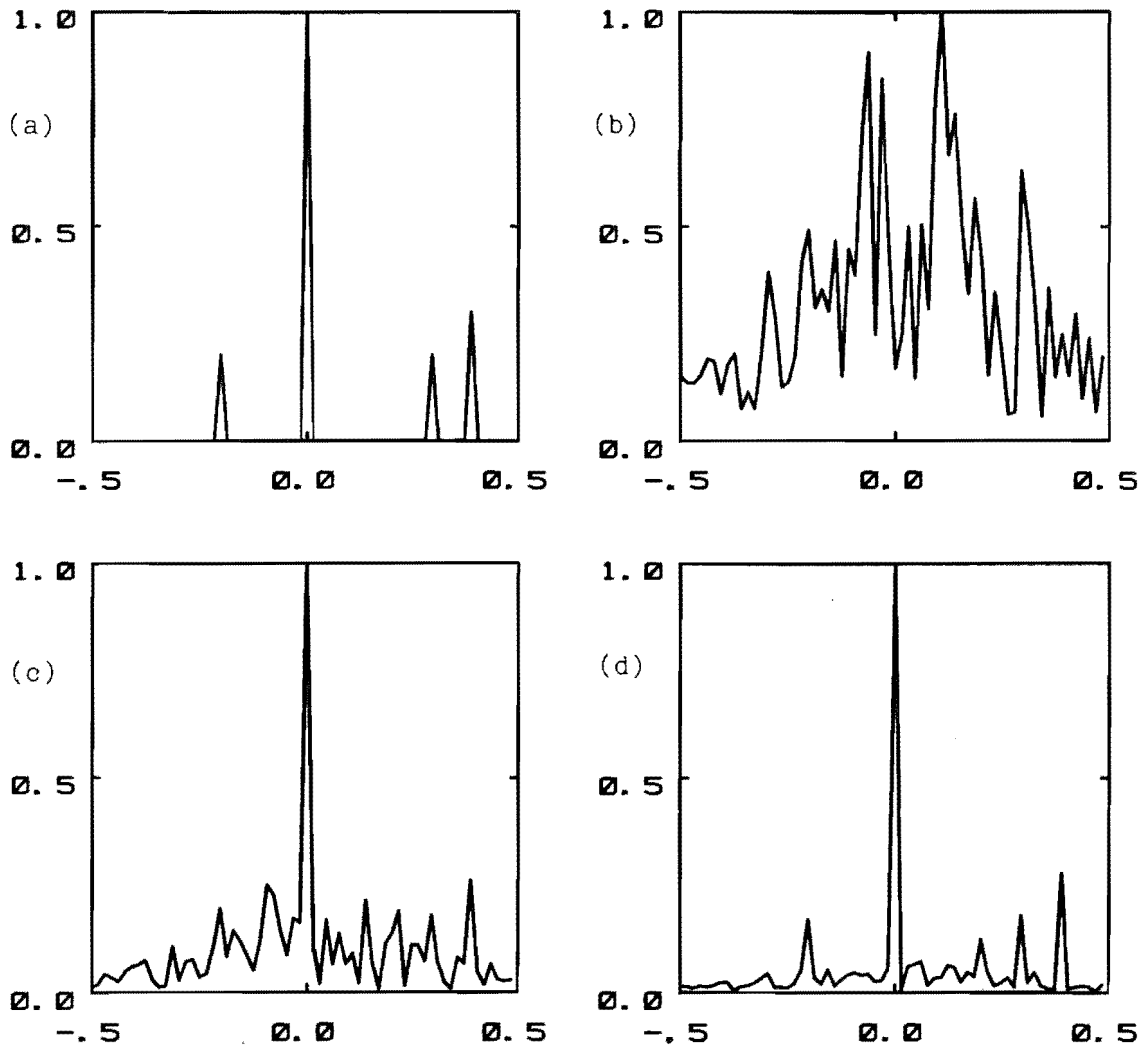


Fig. 6.4 One-dimensional simulation of elementary shift-and-add.
Image magnitudes are displayed.
(a) Object f.
(b) Speckle image s_1.
(c) Elementary shift-and-add image f_e.
16 speckle images averaged.
(d) As for (c) but 128 speckle images averaged.

The objects considered by Robinson (1982) are composed of strands of wire or nylon. To investigate the efficacy of shift-and-add for reconstructing more complicated objects, a target consisting of a metal

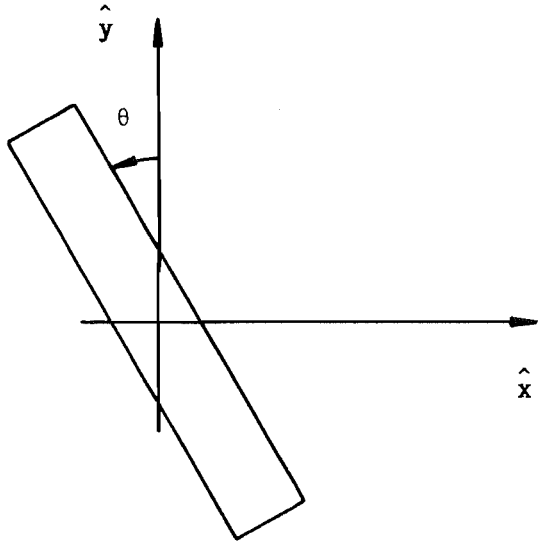


Fig. 6.5 Configuration of strip target. The strip is made of brass and is 5 mm × 0.2 mm. Its orientation is described by θ.

strip was constructed by the cassette method described in ¶5.2.3 (refer to Fig. 6.5). The scattering of ultrasound by the strip was measured using the ultrasonic scanner described in Chapter 5. Four sets of measurements were made with the target at a different orientation for each set. The scattering was obtained for 128 frequency bands, the centre frequencies of which are regularly spaced between 1.7 MHz and 3.4 MHz. 'Narrow-band' images of the target were calculated from the scattering at each band according to the formula (1.242). The weighting function in (1.242) was assumed to be constant over the measurement aperture which is given by (5.22) and (5.23). The effect of this is to alter the shape of the PSF slightly. The narrow-band images were processed by simple shift-and-add and the results are displayed in Fig. 6.6. Invoking the geometrical theory of diffraction (James 1976), the scattering from the strip consists of a field due to the edges and to a specular reflection from the surface of the strip. When the specular reflection does not enter the measurement aperture, only the edges of the strip are visible as in Fig. 6.6 for $\theta = 0°$,

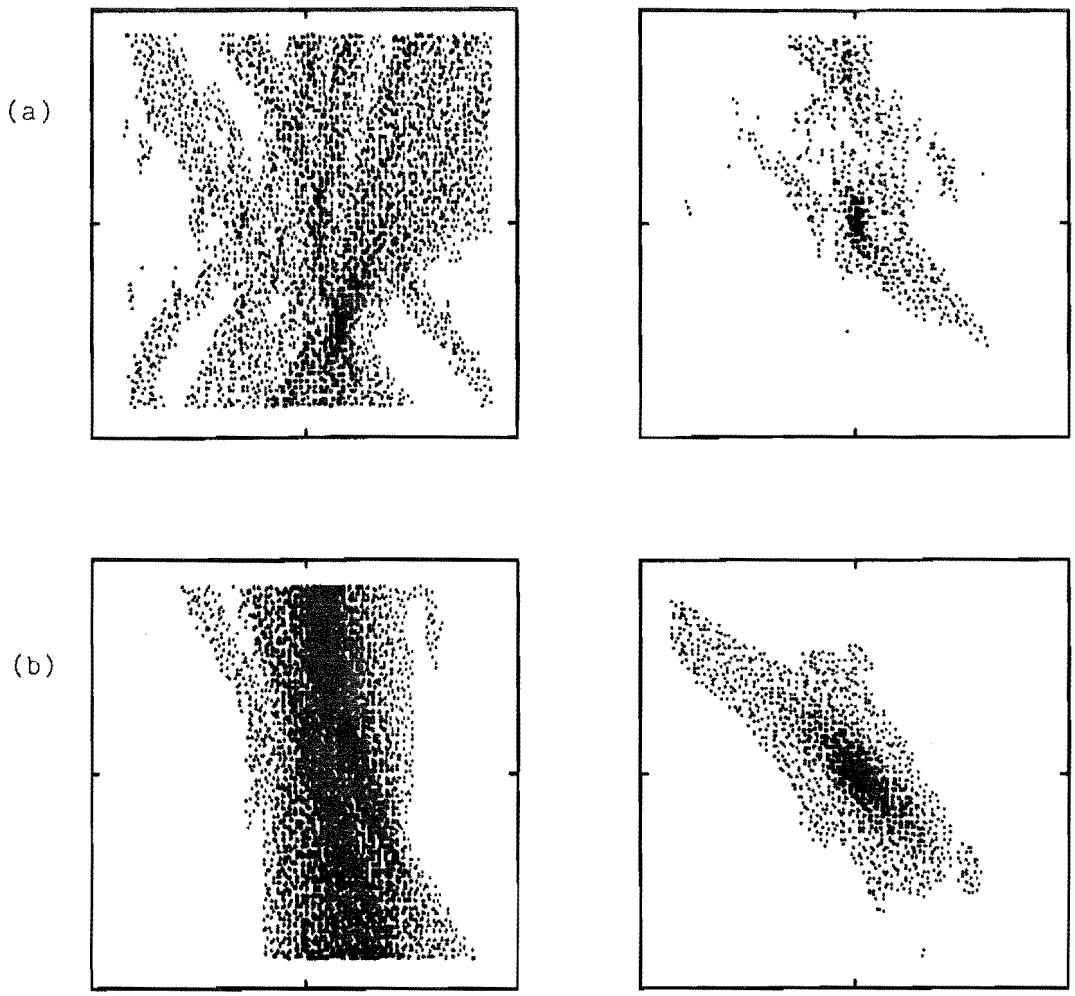


Fig. 6.6 Shift-and-add images of the strip target.
The image coordinates are specified in ¶6.2.
Left column: $|s_{75}|$.
Display parameters: $d_0 = 0$; $d_\varepsilon = 0$; $d_1 = 1.0$.
Right column: $|f_{saa}|$.
Display parameters: $d_0 = -0.07$; $d_\varepsilon = 0$; $d_1 = 1.0$.
(a) $\theta = 0°$. (b) $\theta = 45°$.
(c) $\theta = 90°$. (d) $\theta = 135°$.

90° and 135°. The edges of the strip appear as small (1 mm × 1 mm) regions of increased image intensity. The edge which acts as a reference for shift-and-add appears at (0 mm, 0 mm). The other edge is located at (0 mm, 5 mm) for $\theta = 0°$; (-5 mm, 0 mm) for $\theta = 90°$; and (2.2 mm, 2.2 mm) for $\theta = 45°$. A ghost peak is apparent at (5 mm, 0 mm) for $\theta = 90°$ and at (-2.2 mm, 2.2 mm) for $\theta = 45°$. At $\theta = 45°$, the specular reflection enters

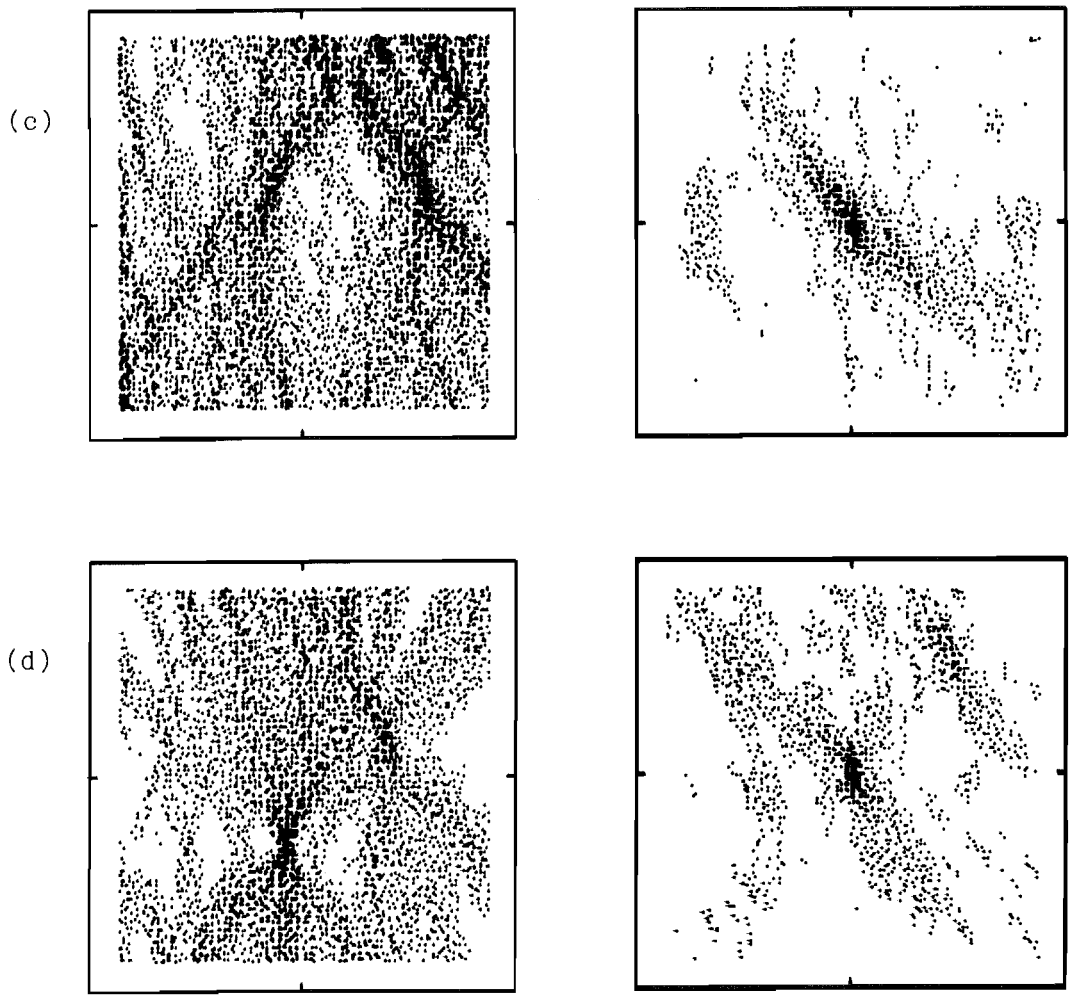


Fig. 6.6 (cont.)

the aperture and the full extent of the strip is visible. The specular reflection swamps the field scattered by the edge. For $\theta = 0°$, $90°$, and $135°$, one of the edges is brighter than the other in all of the narrow band images and acts as the shift-and-add reference. At $\theta = 45°$, each point of the strip is expected to have a comparable magnitude in the narrow-band images. Shift-and-add, therefore selects each point of the strip with approximately equal probability. The shift-and-add image at $\theta = 45°$ is therefore blurred. This is another example of ghosting. The ghost peak in the shift-and-add image for $\theta = 135°$ is caused by high side lobes in the narrow-band PSF (¶1.8.1). The positions of the side lobes depend on frequency. When the narrow-band images are shifted-and-added the side lobes are reduced considerably. Hence, the ghost is expected to be faint and this is indeed the case.

6.5 CORRELATE-AND-ADD

Although shift-and-add is computationally the simplest digital speckle processing technique it was certainly not the first to be discovered. The first was a technique called 'speckle interferometry' proposed by Labeyrie (1970). The power spectra of the speckle images are averaged and corrected by the power spectra obtained from an unresolvable object. The result is the power spectrum of the object, for which the autocorrelation of the object may be obtained. This processing sequence may be written as

$$f * f = \mathcal{F}^{-1}\left\{\frac{\langle |S_m|^2 \rangle_m}{\langle |H_m|^2 \rangle_m}\right\} \tag{6.36}$$

and it works because

$$|S_m|^2 \approx |F|^2 |H_m|^2 \ . \tag{6.37}$$

The capital letter convention defined in ¶1.2.3 for denoting the Fourier transform of a function is invoked in (6.36) and (6.37). Unfortunately, it is not a simple matter to reconstruct an object from its autocorrelation, and it is computationally intensive even in those situations (which actually encompass most two-dimensional images) where the Fourier phase can be recovered from the Fourier magnitude. Bates (1982) reviews speckle processing techniques in optical astronomy. Phase retrieval is reviewed by Bates and Fright (1984). An approach, called 'LWH' processing, which can be regarded as the direct ancestor of the shift-and-add principle, was introduced by Lynds et al. (1976) as an improvement on speckle interferometry in those situations where it permitted images, rather than autocorrelations, to be generated. The method involves expressing each speckle image as a collection of peaks. The positions of the peaks above a certain threshold are assumed to correspond to the locations of replicas of the object. These replicas are then superposed. A more straight-forward procedure (Milner 1979 Ch. 6) which eliminates the need to express the speckle image as a collection of peaks, is to construct a mask function m_k from the speckle image s_k according to

$$m_k = \mathrm{thresh}_d\left(w(s_k), \varepsilon_k\right) \tag{6.38}$$

where w is defined by (6.26) and the discontinuous threshold image operation thresh_d is defined for an arbitrary function g by

$$\text{thresh}_d(g,\varepsilon)(\mathbf{r}) = \begin{cases} g(\mathbf{r}) & |g(\mathbf{r})| \geq \varepsilon \\ 0 & |g(\mathbf{r})| < \varepsilon \end{cases} . \qquad (6.39)$$

In general, the operator defined by (6.39) introduces discontinuities wherever the function g crosses the threshold ε. An alternative continuous threshold operator, $\text{thresh}_{c,n}$ is defined by

$$\text{thresh}_{c,n}(g,\varepsilon)(\mathbf{r}) = g(\mathbf{r})\left[1 - \left(\frac{|g(\mathbf{r})|^{2n}}{\varepsilon^{2n}} + 1\right)^{-1}\right] . \qquad (6.40)$$

For the results presented in this thesis, the discontinuous threshold is employed. The replicas identified by the mask m_k in each speckle image are superposed by cross-correlating the mask with the speckle image. The results of the cross-correlation for each speckle image are added. The result, denoted by f_{caa}, is

$$f_{caa} = \langle m_k * s_k \rangle_k . \qquad (6.41)$$

An appropriate name for this modification of the technique of Lynds et al. (1976) is therefore correlate-and-add. Shift-and-add is similar to correlate-and-add when the ε_k are chosen according to

$$\varepsilon_k = |b(s_k)| . \qquad (6.42)$$

However, if some of the speckle images do not contain a unique bright point the techniques produce different results. For instance, if a speckle image s_k has two bright points $\mathbf{b}_1$ and $\mathbf{b}_2$ which satisfy (6.24), the mask function is

$$m_k = T_{\mathbf{b}_1}(\delta) + T_{\mathbf{b}_1}(\delta) . \qquad (6.43)$$

So two replicas are superposed by correlate-and-add where shift-and-add uses only the one corresponding to $\mathbf{b}(s_k)$. Bagnuolo (1982) defines an average similar to correlate-and-add which he calls Bates(n), where n has a similar role in the average as ε_k has in (6.38).

Some examples of correlate-and-add processing are shown in Fig. 6.7. The object is called 'DD(½)' (the terminology is explained in ¶7.6.3)' and is depicted in Fig. 7.6. In the 'mask-correlated' speckle images $m_k * s_k$, an image of the object can be seen. The image of the object in the f_{caa} images is clearer, however, because 128 mask-correlated speckle images have been averaged. Note that as the mask threshold is reduced, the contamination is reduced. However, the image begins to resemble the autocorrelation of the object. Therefore, a compromise must be reached between obtaining a low level of contamination and a faithful image.

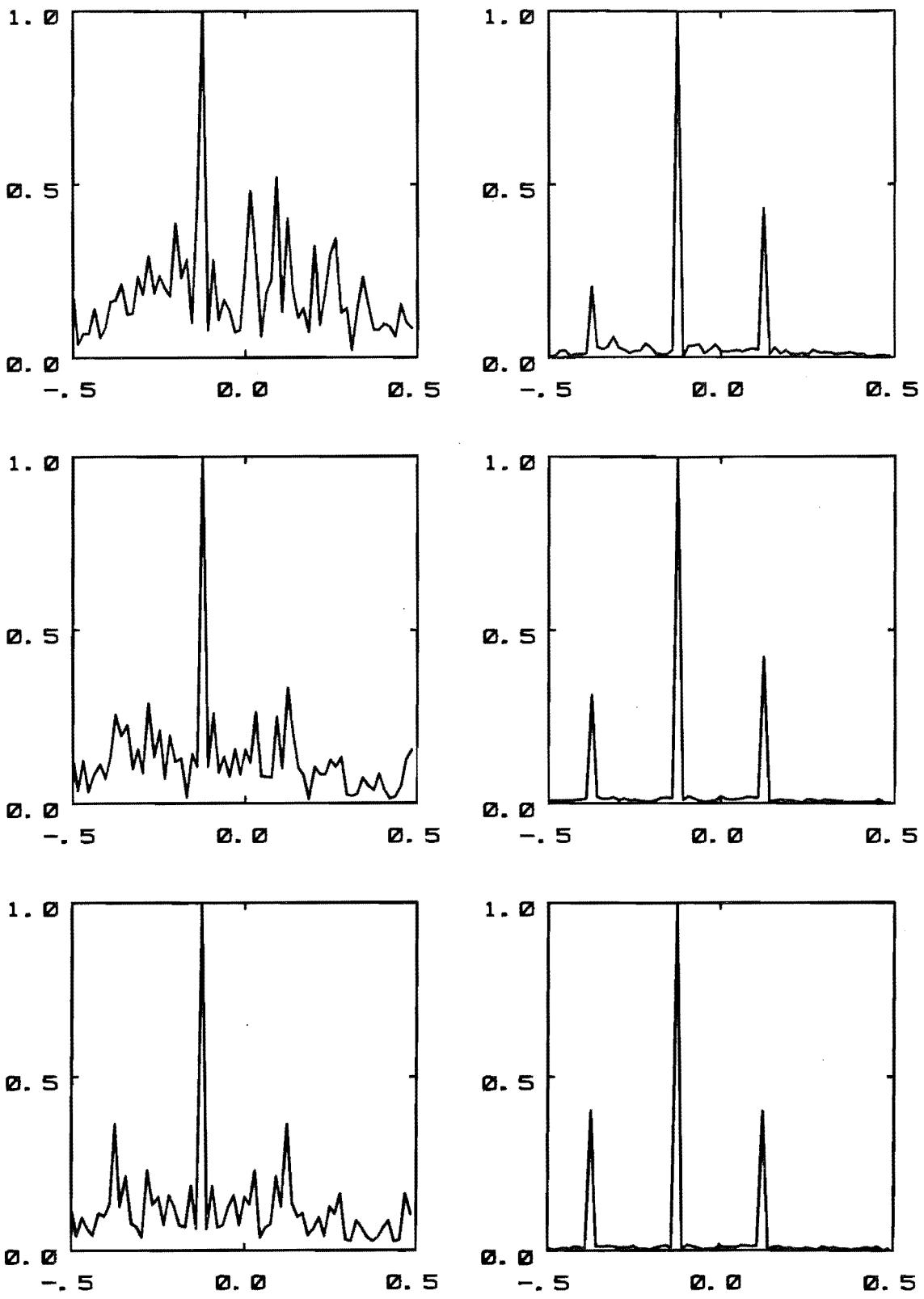


Fig. 6.7 Correlate-and-add simulation. Weighting function has n = 1. The object f is given by

$$f(x) = \delta(x) + \tfrac{1}{2}\delta(x - \tfrac{1}{4}).$$

Left column: $|m_1 * s_1|$. Right column: $|f_{caa}|$.

Top row: $\varepsilon_k = 0.8\ b(s_k)$.

Middle row: $\varepsilon_k = 0.4\ b(s_k)$.

Bottom row: $\varepsilon_k = 0$.

6.6 NATURALLY ARISING REFERENCES

Shift-and-add requires a part of the object to act as a reference. The reference should be bright and of small extent. Such image structure can arise quite naturally. Consider, a smooth surface which reflects ultrasound and assume that it is insonified by a plane wave. If the surface is curved there will be very few directions for which the incident wave is specularly reflected into the aperture of the imaging system. If the surface is quite sharply curved the extent of the region which reflects specularly will also be small. In optics, such a phenomenon is called 'glinting'.

The glinting effect also occurs in ultrasonic imaging as Fig. 6.8 demonstrates. A bovine artery was scanned with and without a wire strand parallel to it. There are four points A, B, C, and D on the cross-section through the artery shown in Fig. 6.8(a) where the incident wave is reflected specularly into the aperture. The y coordinate of the wire relative to the artery is in the range -2 mm to 2 mm. It could not be measured more accurately because of the difficulty in accessing the object region within the tank of the ultrasonic scanner. The wire was intended to act as a reference for shift-and-add but, in fact, a specular reflection from the inner wall of the artery at B is brighter than the wire. The extent of the specular reflection is not much larger than the resolution cell of the imaging system. Hence, it is a suitable reference for shift-and-add. This is shown to be the case by Fig. 6.8(e) because the wire is resolved to the diffraction limit. Note that the image of the wire is not apparent in Fig. 6.8(f) because its intensity is below the display threshold d_ε, and is thus displayed as white. The specular reflections at A and D are apparent in Fig. 6.8(d). The reflection at C is not visible, probably because the incident wave at C is considerably weakened by a strong reflection at E, which by symmetry must be of a similar strength to that at B. The interpretation of the images of the artery in Fig. 6.8 differs from that given by Minard et al. 1985. However, the interpretation presented here accounts for the known dimensions of the target more accurately.

Ultrasound is often used in medicine to guide operating instruments within tissue. Instruments such as needles (Heckman and Seidel 1983) could provide a suitable reference for shift-and-add because the tip scatters ultrasound like a point scatterer. Stern et al. (1983) discusses the construction of scatterers from a planar array of 1/16" diameter stainless steel spheres suitable for biological implantation. While such an object

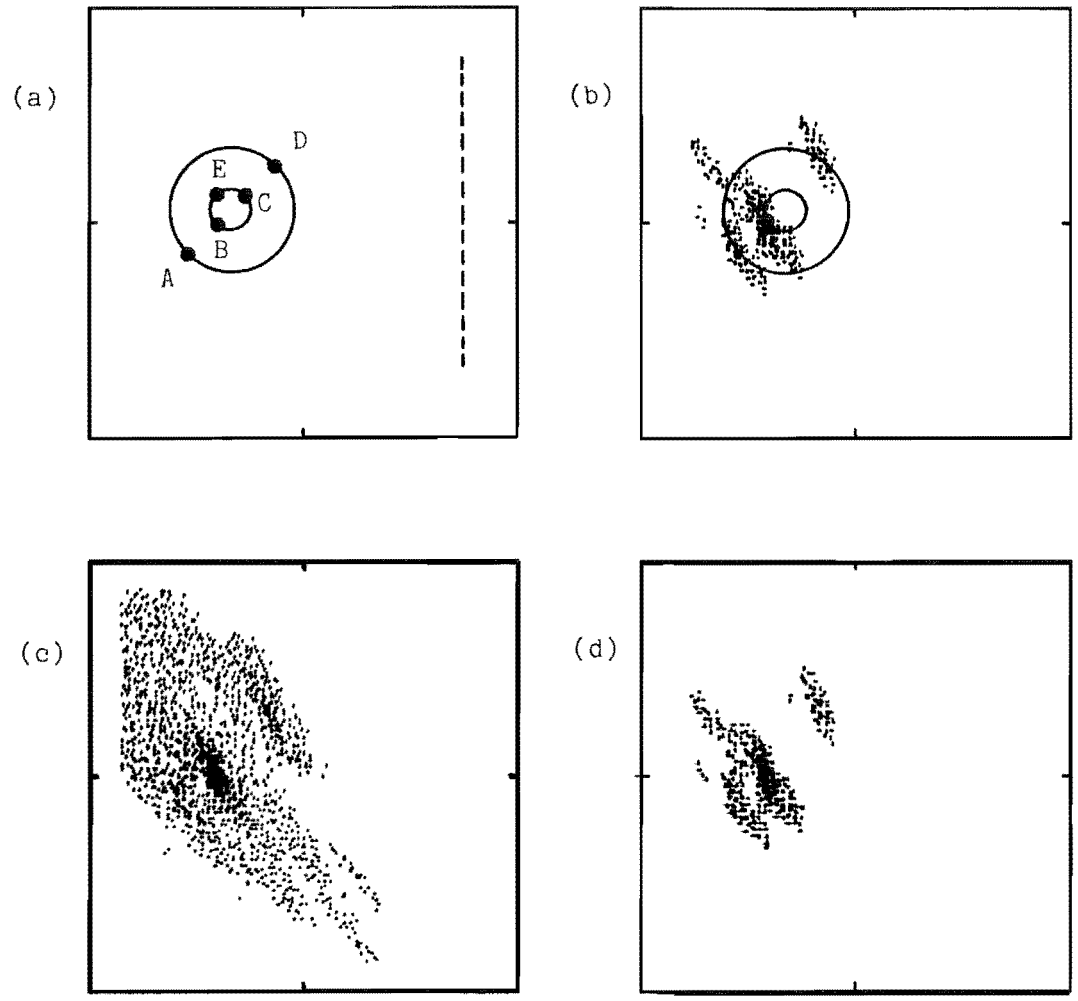


Fig. 6.8 Shift-and-add images of a bovine artery and a copper wire.

Display parameters: $d_0 = -0.07$; $d_1 = 1.0$.

(a) Sketch of cross-section through artery, internal and external diameters 1.3 mm and 4.0 mm. The wire (0.18 mm diameter) is located at a point on the dashed line.

(b) Cross-section of artery superposed with (d).

(c) f_{saa} of artery by itself, $d_\varepsilon = 0$.

(d) As for (c), but $d_\varepsilon = 0.15$.

(e) f_{saa} of artery and wire, $d_\varepsilon = 0$.

(f) As for (e) but $d_\varepsilon = 0.15$.

would not be entirely suitable for ordinary shift-and-add, it could well form a generalised reference for the extended shift-and-add principle

discussed in Chapter 7.

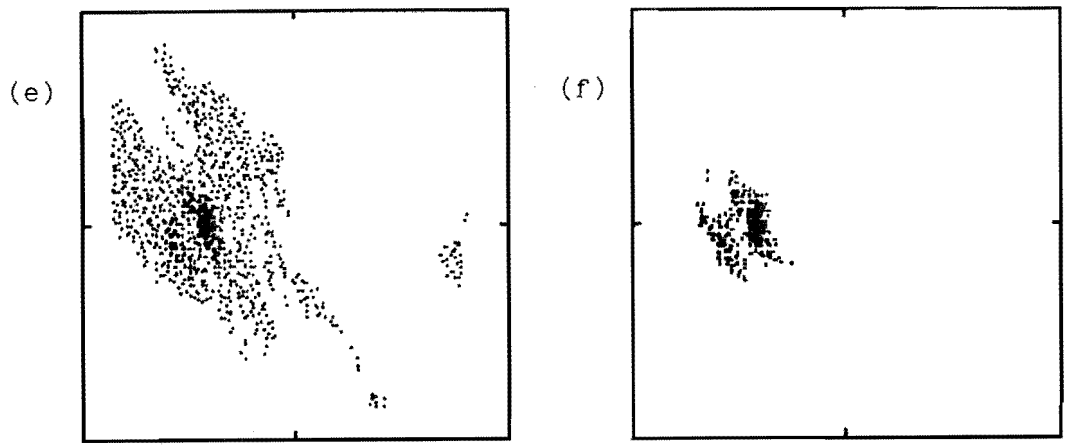


Fig. 6.8 (cont.)

6.7 THE STRUCTURE OF THE WEIGHTED SHIFT-AND-ADD IMAGE

As indicated by the results presented in ¶6.4, shift-and-add produces faithful reconstructions when the object is 'dominated' by a δ function, i.e.

$$f = \delta + g \tag{6.44}$$

where g is fainter than the δ function in some sense, e.g.

$$\int_{T_{\mathbf{r}}} |g(\mathbf{r})|^2 \, d\mathbf{r} \ll 1 \ . \tag{6.45}$$

When g is not small an artefact called 'ghosting' is significant. Ghosting is evident in the shift-and-add images of the strip object oriented at $\theta = 90°$ shown in ¶6.4. An extra peak or 'ghost' tends to appear in f_{saa}. It arises because by chance a subpeak is brighter than the peak in some of the speckle images and thus the replica is shifted incorrectly. A simple heuristic model is proposed in ¶6.7.1 for this artefact. The only images which are devoid of ghosts, i.e. are 'ghostless', are those of a single δ function.

The ghosting artefact causes the shift-and-add image to have twice the extent of the object. This may be explained in one dimension with reference to Fig. 6.9. Let the position vectors of points at both edges of the object be at $\mathbf{r}_1$ and $\mathbf{r}_2$. When the pixel at $\mathbf{r}_2$ is mistaken for the

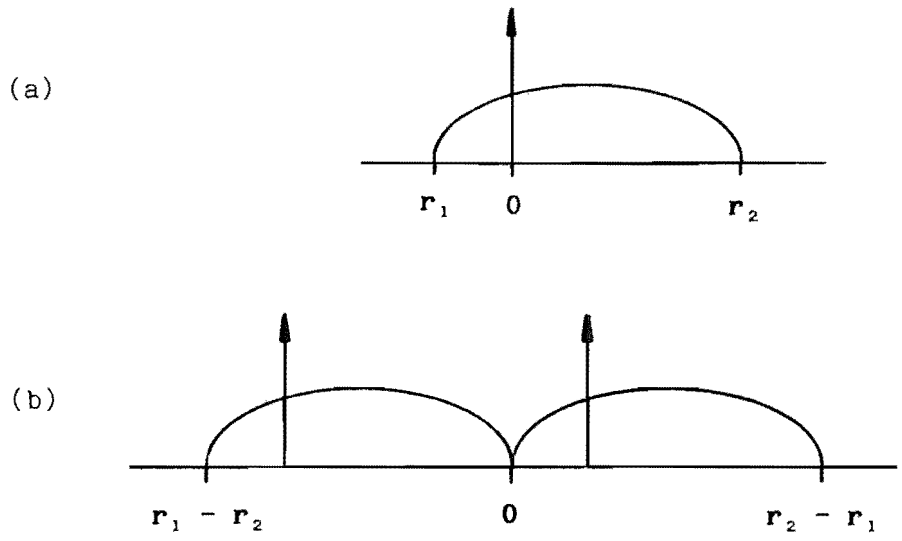


Fig. 6.9 The ghosting artefact is responsible for the shift-and-add image having twice the extent of the object. When the left edge of the object shown in (a) is mistaken for the reference, the object is shifted to the rightmost of the two positions shown in (b) and vice versa.

reference it is shifted to the origin. The point $\mathbf{r}_1$ is shifted to $\mathbf{r}_1 - \mathbf{r}_2$. Vice versa, when the pixel at $\mathbf{r}_1$ is mistaken for the reference, $\mathbf{r}_2$ is shifted to $\mathbf{r}_2 - \mathbf{r}_1$. Hence, the extent of the shift-and-add image is $(\mathbf{r}_2 - \mathbf{r}_1) - (\mathbf{r}_1 - \mathbf{r}_2) = 2(\mathbf{r}_2 - \mathbf{r}_1)$ which is twice the extent of the object.

6.7.1 A General Object

A finite resolution model similar to that employed for the speckle PSF (6.7) is here used to describe a general object,

$$f_k = \sum_k f_k \, T_{\mathbf{f}_k}(\delta) \tag{6.46}$$

where $\{f_k\}$ is a set of complex amplitudes and $\{\mathbf{f}_k\}$ is a set of position vectors. A speckle image of f for the speckle PSF given by (6.7) is found by substituting (6.7) and (6.46) into (6.17),

$$s = \sum_l \sum_k h_l \, f_k \, T_{\mathbf{h}_l + \mathbf{f}_k}(\delta) \, . \tag{6.47}$$

The brightest point of s must be one of the $\mathbf{h}_l + \mathbf{f}_k$, say

$$\mathbf{b}(s) = \mathbf{h}_{l_1} + \mathbf{f}_{k_1} \, . \tag{6.48}$$

If f has a reference and $\mathbf{f}_{k_1}$ is not the location of the reference then ghosting occurs. The probability that the brightest point corresponds to $\mathbf{f}_{k_1}$, irrespective of the value of $\mathbf{h}_{l_1}$ is denoted by p_{k_1}. The most probable speckle image amplitude at $\mathbf{b}(s)$ is

$$b(s) = h_{l_1} f_{k_1} \, . \tag{6.49}$$

The possibility that another of the δ functions in (6.47) overlaps with $\mathbf{b}(s)$, i.e. there exists l_2 and k_2 such that

$$\mathbf{h}_{l_1} + \mathbf{f}_{k_1} = \mathbf{h}_{l_2} + \mathbf{f}_{k_2} \, , \tag{6.50}$$

is ignored in (6.49). Since b(s) is the maximum value of s it is likely that

$$|h_{l_1}| \approx |h_L| \tag{6.51}$$

where h_L is the largest of the $\{h_L\}$. Without loss of generality, the expected value of h_{l_1} is assumed to be unity. Hence, the weighted and shifted speckle image is given by

$$w(b(s)) \, T_{-\mathbf{b}(s)}(s) = |f_{k_1}|^n \, T_{-\mathbf{f}_{k_1}}(f) + c' \tag{6.52}$$

where c' is a random contamination of δ functions. The index k_1 is now replaced with k. The expected value of n-weighted shift-and-add is found by using the probability $\{p_k\}$ to be

$$f^{(n)}_{wsaa} = \sum_k p_k \, |f_k|^n \, T_{-\mathbf{f}_k}(f) \, . \tag{6.53}$$

Now

$$T_{-\mathbf{f}_k}(f) = T_{-\mathbf{f}_k}(\delta * f) = T_{\mathbf{f}_k}(\delta) * f \tag{6.54}$$

where (1.46) and the sifting property of the δ function are invoked. Hence, the n-weighted shift-and-add image is given by

$$f^{(n)}_{wsaa} = p(f) * f \tag{6.55}$$

where p(f) is defined by

$$p(f) = \sum_k p_k \, |f_k|^n \, T_{\mathbf{f}_k}(\delta) \; . \tag{6.56}$$

Thus (6.55) shows that the n-weighted shift-and-add image is the cross-correlation of an object-dependent 'ghosting' function p(f) with the object.

6.7.2 An Object With a Reference.

The function p(f) has a special form when the object is given by (6.44) and is dominated by the reference. Let k = 0 in (6.46) correspond to the reference and denote the remaining f_k and $\mathbf{f}_k$ by g_k and $\mathbf{g}_k$ respectively. Since the reference is strong and of unit amplitude (6.34),

$$|g_k| \ll 1 \; . \tag{6.57}$$

The function g is given by

$$g = \sum_k g_k \, T_{\mathbf{g}_k}(\delta) \; . \tag{6.58}$$

Firstly, consider the structure of the speckle image, which is obtained by substituting (6.58) and (6.44) into (6.17),

$$s = h + \sum_k g_k \, T_{\mathbf{g}_k}(h) \; . \tag{6.59}$$

This illustrated in Fig. 6.10. The speckle image consists of replicas of the speckle PSF located at the positions of each component of the object. This is an alternative viewpoint to that expressed by (6.18). In the right hand side of (6.59), h is called the 'reference' replica and $g_k \, T_{\mathbf{g}_k}(h)$ is called the 'g_k-replica'. Upon substituting (6.9) into (6.59), the formula

$$s = \sum_l h_l \, T_{\mathbf{h}_l}(\delta) + \sum_l \sum_k h_l \, g_k \, T_{\mathbf{h}_l + \mathbf{g}_k}(\delta) \tag{6.60}$$

is obtained. This is similar to (6.47) but now the role of the reference can be seen more clearly. When ghosting occurs the brightest point occurs in a replica of the PSF other than the reference replica. In other words, for some l_1 and k_1

$$b(s) = \mathbf{h}_{l_1} + \mathbf{g}_{k_1} \; . \tag{6.61}$$

From Fig. 6.10 and since $|g_k| \ll 1$ it is not obvious how the situation represented by (6.61) can occur. However, it should be noted that the replicas shown in Fig. 6.10 in fact overlap and that a δ function at $\mathbf{h}_{l_2}$, belonging to the reference replica, has reinforced the δ function at $\mathbf{h}_{l_1} + \mathbf{g}_{k_1}$, belonging to the g_k-replica, i.e.

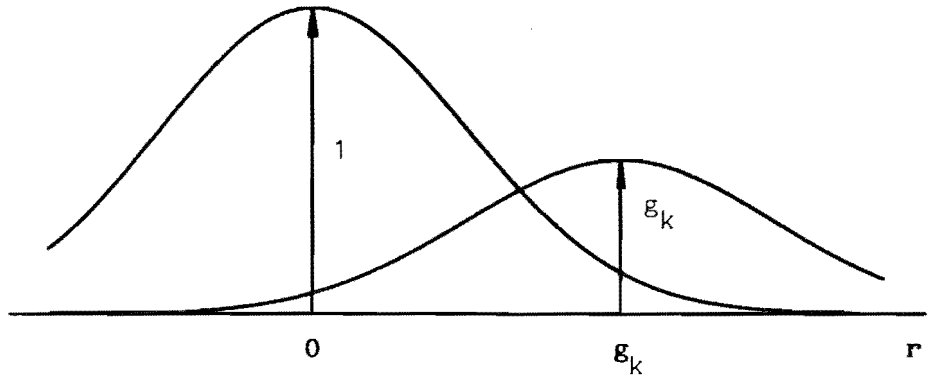


Fig. 6.10 A speckle image of an object with a reference. The image consists of replicas of the speckle PSF at the location of the reference and at every other component of the object. A replica of the PSF is portrayed by the seeing disc $o(\mathbf{r})$.

$$\mathbf{h}_{l_2} = \mathbf{h}_{l_1} + \mathbf{g}_{k_1} \tag{6.62}$$

and

$$|h_{l_2} + g_{k_1} h_{l_1}| > |h_L| \ . \tag{6.63}$$

Since (6.63) refers to g_{k_1} and not to g_k for $k \neq k_1$, it is reasonable to assume that p_{k_1} depends only on g_{k_1} and not g_k for $k \neq k_1$. A similar argument for (6.62) shows that p_{k_1} does not depend on $\mathbf{g}_k$ for $k \neq k_1$. Hence p_{k_1} may be written

$$p_{k_1} = \mu(g_{k_1}, \mathbf{g}_{k_1}) \tag{6.64}$$

where the function μ is called the 'ghosting potential'. Since the probabilities p_k satisfy

$$\sum_k p_k = 1 \ , \tag{6.65}$$

the probability p_0 of selecting a δ function in the reference replica is given by

$$p_0 = 1 - \sum_k \mu(g_k, \mathbf{g}_k) \ . \tag{6.66}$$

The ghosting function is expressed in terms of the ghosting potential by substituting (6.66) and (6.64) into (6.56), and invoking (6.58), i.e.

$$p(f) = p_o \, |f_o|^n \, T_{\mathbf{f}_o}(\delta) + \sum_k p_k \, |g_k|^n \, T_{\mathbf{g}_k}(\delta)$$

$$= \left[1 - \sum_k \mu(g_k, \mathbf{g}_k)\right] \delta + \mu(g) \, |g|^n$$

$$= \left[1 - \int_{T_\mathbf{r}} \mu(g)\right] \delta + \mu(g) \, |g|^n \qquad (6.67)$$

where $\mu(g)$ is defined by

$$\mu(g) = \sum_k \mu(g_k, \mathbf{g}_k) \, T_{\mathbf{g}_k}(\delta) \, . \qquad (6.68)$$

Substituting (6.67) into (6.55), one obtains

$$f^{(n)}_{wsaa} = \left[1 - \int_{T_\mathbf{r}} \mu(g)\right] f + \left[\mu(g) \, |g|^n\right] * f \, . \qquad (6.69)$$

The effect of ghosting on the n-weighted shift-and-add image can be clearly seen from (6.69). The first term on the right hand side of (6.69) is the reconstruction of the object. When the integral of $\mu(g)$ is significant with respect to unity (the integral of the reference amplitude) the second term is a significant part of $f^{(n)}_{wsaa}$. This term is responsible for the ghosting artefact.

6.7.3 The Ghosting Potential

A formula for the ghosting potential is now derived. To simplify the notation the g_k and $\mathbf{g}_k$ are denoted by α and $\boldsymbol{\alpha}$ respectively. When a δ function at $\mathbf{h}_1$ from the reference replica is chosen, the speckle image amplitude is given by (refer to Fig. 6.10)

$$s(\mathbf{h}_1) = h_1 + \alpha h_j \qquad (6.70)$$

where the possibility that the α-replica contributes is included. Denote the magnitude of $s(\mathbf{h}_1)$ by A_o,

$$A_o = |s(\mathbf{h}_1)| \, . \qquad (6.71)$$

It is now assumed that the h_1 are bounded by the seeing disc o and that A_o is a uniform random variable. The brightest point is likely to come from the centre of the replica, hence

$$|h_1| \approx 1 \qquad (6.72)$$

and therefore

$$1 - |\alpha|\sigma(\boldsymbol{\alpha}) \leq A_0 \leq 1 + |\alpha|\sigma(\boldsymbol{\alpha}) \ . \tag{6.73}$$

In actuality, the histograms of ¶6.3.2 show that A_0 is more closely modelled by a Rayleigh distribution (Goodman 1975), however, (6.73) is a reasonable first order approximation. Now consider the case when the brightest point lies in the α-replica. The equations corresponding to (6.70) to (6.73) are

$$s(\boldsymbol{\alpha} + \mathbf{h}_1) = h_j + \alpha h_1 \ , \tag{6.74}$$

$$A_\alpha = |s(\boldsymbol{\alpha} + \mathbf{h}_1)| \ , \tag{6.75}$$

and

$$|\alpha| - \sigma(\boldsymbol{\alpha}) \leq A_\alpha \leq |\alpha| + \sigma(\boldsymbol{\alpha}) \ . \tag{6.76}$$

The probability $\mu(\alpha,\boldsymbol{\alpha})$ that the brightest point lies in the α-replica rather than the reference replica is the probability that $A_\alpha > A_0$. The possible values for A_0 and A_α are illustrated by the rectangle ABCD in Fig. 6.11. First consider the case when there exist values for which $A_\alpha > A_0$. These values are indicated by the triangle DEF. Since the random variables A_0 and A_α are assumed to be uniformly distributed,

$$\mu(\alpha,\boldsymbol{\alpha}) = \frac{\text{area of triangle DEF}}{\text{area of rectangle ABCD}} \ . \tag{6.77}$$

For other than uniform distributions, a multiple integration of the appropriate probability density functions could be employed. From inspection of Fig. 6.11,

$$\mu(\alpha,\boldsymbol{\alpha}) = \tfrac{1}{2}\frac{[(|\alpha|+\sigma(\boldsymbol{\alpha})) - (1 - |\alpha|\sigma(\boldsymbol{\alpha}))]^2}{2|\alpha|\sigma(\boldsymbol{\alpha}) \times 2\sigma(\boldsymbol{\alpha})} \ . \tag{6.78}$$

In the second case when there are no values for which $A_\alpha > A_0$, $\mu(\alpha,\boldsymbol{\alpha}) = 0$. Hence,

$$\mu(\alpha,\boldsymbol{\alpha}) = \begin{cases} \tfrac{1}{8}\dfrac{[|\alpha|\sigma(\boldsymbol{\alpha}) + |\alpha| + \sigma(\boldsymbol{\alpha}) - 1]^2}{|\alpha|\sigma(\boldsymbol{\alpha})^2} & |\alpha| > \dfrac{1-\sigma}{1+\sigma} \\ 0 & |\alpha| \leq \dfrac{1-\sigma}{1+\sigma} \ . \end{cases} \tag{6.79}$$

The dependence of $\mu(g)$ upon the object $\delta + g$ and the seeing disc σ is given by (6.79). The form of (6.79) is more easily understood for certain special cases, which are now investigated.

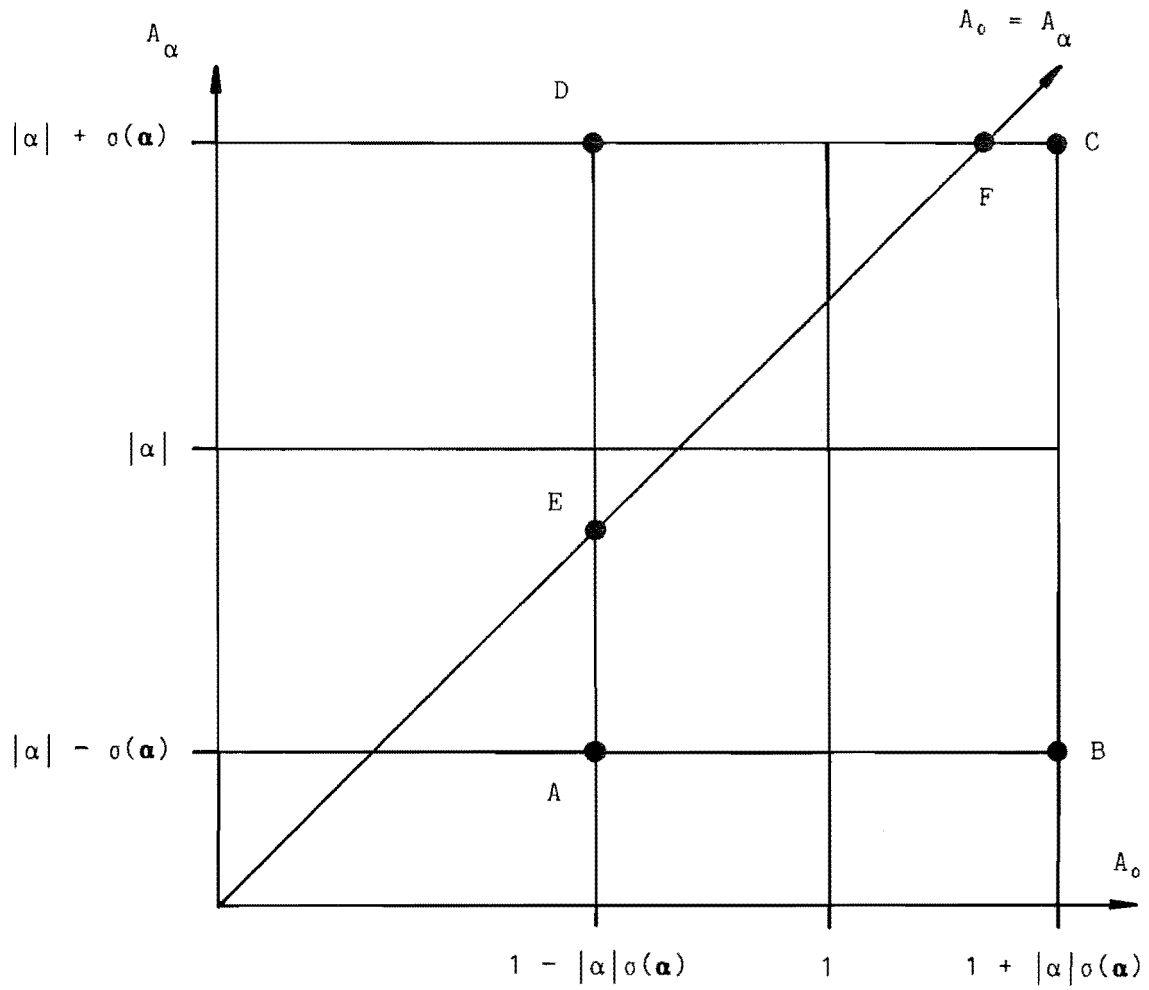


Fig. 6.11 The ghosting potential. The possibilities $A_\alpha > A_o$ are shown as the triangle DEF. All of the values which A_o and A_α may take are shown by the rectangle ABCD.

6.7.4 A Gaussian Seeing Disc

Consider the case of a Gaussian seeing disc,

$$o(\mathbf{r}) = \exp(-\pi|\mathbf{r}|^2/d^2) \tag{6.80}$$

and define a normalised position vector $\hat{\boldsymbol{\alpha}}$ by

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha}/d \ . \tag{6.81}$$

For a position near the centre of the seeing disc, the ghosting potential

is obtained by applying Taylor's series expansion to (6.79),

$$\mu(\alpha,\boldsymbol{\alpha}) = \frac{|\alpha|}{2}\left[1 - \left(\frac{1}{|\alpha|} - 1\right)\pi|\hat{\boldsymbol{\alpha}}|^2\right] , \quad |\hat{\boldsymbol{\alpha}}|^2 << 1 . \tag{6.82}$$

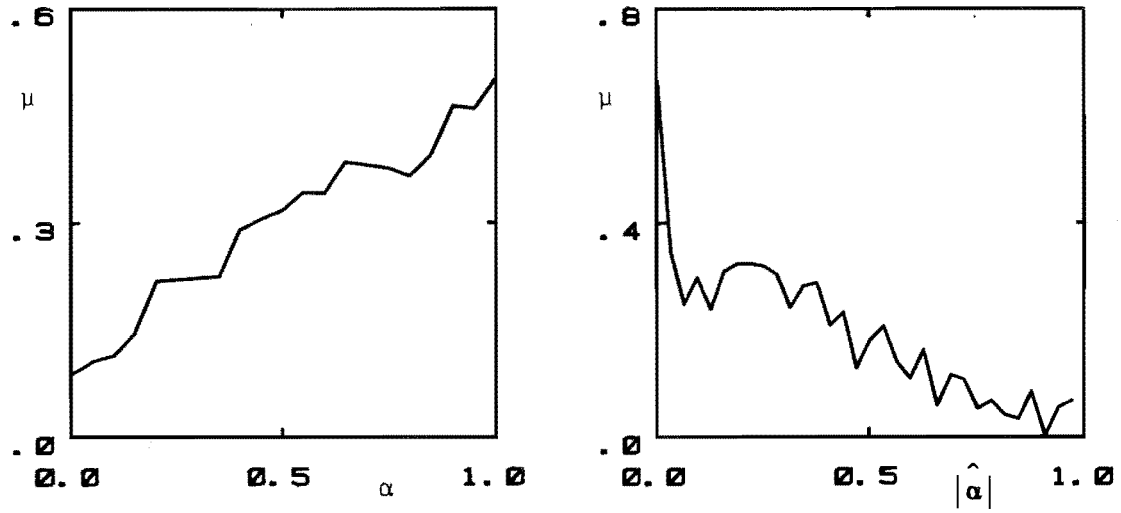


Fig. 6.12 The ghosting potential μ.
(a) $\mu(\alpha, \hat{\boldsymbol{\alpha}} = \frac{1}{8}\hat{\mathbf{x}})$ as a function of α.
(b) $\mu(\alpha = \frac{1}{2}, \hat{\boldsymbol{\alpha}})$ as a function of $|\hat{\boldsymbol{\alpha}}|$.

The correctness of this formula is supported by the simulation results shown in Fig. 6.12. The form of $\mu(\alpha,\boldsymbol{\alpha})$ for $\hat{\boldsymbol{\alpha}} = \frac{1}{8}\hat{\mathbf{x}}$ was obtained by simulation and is shown in Fig. 6.12(a). This should be compared with the theoretical prediction for this case, from (6.82) of

$$\mu(\alpha, \hat{\boldsymbol{\alpha}} = \tfrac{1}{8}\hat{\mathbf{x}}) = \tfrac{1}{2}\alpha . \tag{6.83}$$

The simulated result is linear with a slope of 0.49 and an offset of 0.09 in good agreement with (6.83). Also, the form of $\mu(\alpha,\boldsymbol{\alpha})$ for $\alpha = \frac{1}{2}$ was obtained by simulation (Fig. 6.12(b)). The theoretical result for this case is

$$\mu(\tfrac{1}{2}, \hat{\boldsymbol{\alpha}}) = \tfrac{1}{4}[1 - \pi|\hat{\boldsymbol{\alpha}}|^2] . \tag{6.84}$$

The principal difference between the theoretical and simulated results is a spike at $\hat{\boldsymbol{\alpha}} = 0$. This may be due to the way in which $\mu(\alpha,\boldsymbol{\alpha})$ is deduced from the shift-and-add image, or it may be due to the failure of the finite resolution model of the object. The simulated result is otherwise in

reasonable agreement with the theoretical result. It has the form of a negative parabola intersecting the y axis at 0.3 and the x axis at 0.9. Theoretically, it should intersect the y axis at 0.25 and the x axis at 0.6 (from (6.79)). The poor agreement between the x axis intercepts is probably due to the approximations invoked in this subsection being less accurate at the edge of the seeing disc. Nevertheless the overall agreement between theory and simulation is remarkable considering the simplicity of the model employed.

6.7.5 Wide Seeing Disc

Consider the case for which the seeing disc is much wider than the object, i.e.

$$|\mathbf{r}| \ll d \tag{6.85}$$

and

$$o(\mathbf{r}) \approx 1 \ . \tag{6.86}$$

The form of the ghosting potential is very simple for this case. When (6.86) is substituted into (6.79) one obtains

$$\mu(g) = \tfrac{1}{2}|g| \ . \tag{6.87}$$

The ghosting function for this case is found by substituting (6.87) into (6.67) to give

$$p(f) = \left[1 - \tfrac{1}{2}\int_{T_{\mathbf{r}}} |g|\right] \delta + \tfrac{1}{2} g|g|^n \ . \tag{6.88}$$

For elementary shift-and-add, i.e. n = 0, (6.44) is invoked to write (6.88) as

$$p(f) = \tfrac{1}{2}\left[1 - \int_{T_{\mathbf{r}}} |g|\right] \delta + \tfrac{1}{2} f \ . \tag{6.89}$$

Finally, an expression for the elementary shift-and-add image f_e is obtained by substituting (6.89) into (6.55),

$$f = \tfrac{1}{2}\left[1 - \int_{T_{\mathbf{r}}} |g|\right] f + \tfrac{1}{2}|f| * f \ . \tag{6.90}$$

In addition to the comments made following the derivation of (6.69), it is now observed that the ghosting artefact approximates the autocorrelation of the object. Furthermore, the level of the ghosting artefact is determined by the integral of $|g|$ rather than $|g|^2$, as is suggested by (6.45). However, note that (6.90) is derived under the assumption that $|g| \ll 1$,

and may not therefore be applicable in general.

Bagnuolo (1982) and Hunt et al. (1983) analyse shift-and-add on incoherent speckle images. The method described in this section has the advantage of providing more physical insight into how ghosting occurs and what factors govern how it affects the shift-and-add image.

6.8 AN APPROXIMATE SOLUTION TO THE REMOTE PROBING PROBLEM

The Fourier imaging solution to the remote probing problem derived in ¶1.7 assumes that either the Born or the Rytov approximation is valid and that the measurements are free from noise. In practice, inhomogeneities in the reference medium and the limitations of the apparatus mean that measured data differs from that expected within the Born or Rytov approximations. Most successful solutions to the remote probing problem are very simple, for instance the echo-location principle described in ¶2.2.1. There is a large gap between the methods of Fourier imaging, which ignore multiple diffraction and refraction, and full-wave theories (Chadan and Sabatier 1977; Bates 1984). The latter are much too complicated to form the basis of a practical imaging system. In this section it is shown that shift-and-add may be combined with Fourier imaging to produce a solution which is at least partially insensitive to inhomogeneities in the reference medium and to measurement aberrations.

The case of Fourier imaging within the Born approximation using 'right' scattering measurements (¶1.8.5) is considered. The effect of the imperfections mentioned above can be modelled by assuming that the measured data χ' are related to the ideal data χ, which would have been obtained in the absence of imperfections, by

$$\chi'(u|k) = H(u|k)\ \chi(u|k) + C(u|k)\ . \tag{6.91}$$

$H(u|k)$ represents those parts of the imperfections which are the same for all scatterers within the object, i.e. the 'isoplanatic' component. $C(u|k)$ represents measurement noise and the non-isoplanatic component. A diffraction-limited (¶1.8.1) image $\hat{q}(\mathbf{r}|k)$ of the object is obtained from the ideal data using the formula (see (1.242))

$$\hat{q}(\mathbf{r}|k) = [\exp(jk\eta)/2\pi]\ L\{\chi\}(\mathbf{r}|k) \tag{6.92}$$

where

$$L\{\chi\}(\mathbf{r}|k) = \mathcal{F}_u\left\{w(u)\ \chi(u|k)\ \exp\left[-jk\xi(1 - \lambda^2u^2)^{1/2}\right]\right\}\ . \tag{6.93}$$

The components of the position vector $\mathbf{r}$ are (ξ,η) in the coordinate system shown in Fig. 1.5, and the notation of Chapter 1 is invoked. The operator L describes the imaging operation. When it is applied to the measured data an image $\hat{q}'(\mathbf{r}|k)$ is obtained,

$$\hat{q}'(\mathbf{r}|k) = [\exp(jk\eta)/2\pi]\, L\{\chi'\}(\mathbf{r}|k) \ . \tag{6.94}$$

This image can be related to $\hat{q}$ by a convolution. Note that because of the one-dimensional Fourier transform in L,

$$\begin{aligned} L\{\chi'\} &= L\{H\chi + C\} \\ &= L\{H\} \circledcirc_\eta L\{\chi\} + L\{C\} \\ &= h \circledcirc_\eta \hat{q} + c \end{aligned} \tag{6.95}$$

where $\circledcirc_\eta$ denotes a one-dimensional convolution with respect to η, the PSF h is defined by

$$h = L\{H\}, \tag{6.96}$$

and the contamination c is defined by

$$c = L\{C\} \ . \tag{6.97}$$

Thus $\hat{q}'$ is related to $\hat{q}$ by

$$\hat{q}' = [\exp(jk\eta)/2\pi]\, (h \circledcirc_\eta \hat{q} + c) \ . \tag{6.98}$$

Now consider the result of applying elementary shift-and-add (¶6.4.4) to an ensemble $\{\hat{q}'(\mathbf{r}|k_m)\}$ of measured monochromatic images obtained at a set of wave numbers $\{k_m\}$. The wave numbers of which there are M are uniformly distributed between k_{min} and k_{max}, i.e.

$$k_m = k_{min} + (m-1)\Delta k \ , \quad m = 1, \ldots, M \tag{6.99}$$

where

$$\Delta k = (k_{max} - k_{min})/M \ . \tag{6.100}$$

Assume the finite resolution model for the function h,

$$h(\mathbf{r}|k_m) = \sum_l h_{m,l}\, T_{\mathbf{h}_{m,l}}(\delta) \ . \tag{6.101}$$

The one-dimensional convolution in (6.98) is then given by

$$h \circledcirc_\eta \hat{q} = \sum_l h_{m,l}\, T_{\mathbf{g}_{m,l}}(\hat{q}) \tag{6.102}$$

where the position vectors $\mathbf{g}_{m,l}$ are related to the $\mathbf{h}_{m,l}$ by

$$\mathbf{g}_{m,l} = \mathbf{h}_{m,l} - \mathbf{h}_{m,l} \cdot \hat{\xi} \tag{6.103}$$

and therefore have no component in the $\hat{\xi}$ direction. If the object image $\hat{q}(\mathbf{r})$, and hence $\hat{q}(\mathbf{r}|k)$, has a strong unit reference at the origin, then the brightest point of $\hat{q}'(\mathbf{r}|k_m)$ is given by

$$\mathbf{b}\left(\hat{q}'(\mathbf{r}|k_m)\right) = \mathbf{g}_{m,L(m)} \tag{6.104}$$

where L(m) is given by (6.19) and the amplitude of the brightest point is

$$b\left((\hat{q}'(\mathbf{r}|k_m)\right) = [\exp(jk_m\eta)/2\pi]\, h_{m,L(m)} \ . \tag{6.105}$$

The elementary shift-and-add image is defined by (6.27), (6.28), and (6.29), for n = 0, to be

$$\begin{aligned} f_e(\mathbf{r}) &= \frac{1}{M}\, 2\pi \sum_{m=1}^{M} \frac{T_{-\mathbf{g}_{m,L(m)}}\left[\hat{q}'(\mathbf{r}|k_m)\right]}{\exp(jk_m\eta)\, h_{m,L(m)}} \\ &= \frac{1}{M} \sum_{m=1}^{M} \left[\hat{q}(\mathbf{r}|k_m) + c_q(\mathbf{r}|k_m)\right] \\ &= \frac{1}{k_{min} - k_{max}} \int_{k_{min}}^{k_{max}} \hat{q}(\mathbf{r}|k)\, dk \\ &= \frac{1}{k_{min} - k_{max}}\, \hat{q}(\mathbf{r}) \ , \end{aligned} \tag{6.106}$$

where c_q is a combination of randomly shifted replicas of $\hat{q}(\mathbf{r}|k)$, the contamination is assumed to sum to zero, and (1.238) is invoked. The result (6.106) is very significant because it shows that elementary shift-and-add produces 'clean' images (Bates et al. 1983) of the object.

6.8.1 Multiple Scattering

Shift-and-add also reduces the effect of multiple scattering. The induced sources s within the object are obtained from (1.140),

$$s(\mathbf{r}) = q(\mathbf{r})\, \psi(\mathbf{r}) \ . \tag{6.107}$$

The induced sources are due to excitation of the scatterer $q(\mathbf{r})\delta(\mathbf{r})$ by the incident wave $\psi_{gen}(\mathbf{r})$ and by multiple scattering $A(\mathbf{r})$ from other scatterers within the object,

$$\psi(\mathbf{r}) = \psi_{gen}(\mathbf{r}) + A(\mathbf{r}) \ . \tag{6.108}$$

The image reconstruction algorithm L produces an image of the induced

sources, for the plane wave insonification given by (1.180) and with multiple scattering neglected,

$$L\{\chi\} = 2\pi\exp(-jk\eta)\ \hat{q} \ . \tag{6.109}$$

When multiple scattering is considered, (6.109) is modified to

$$L\{\chi\} = 2\pi s/A_{gen} \ . \tag{6.110}$$

By comparing (6.110) and (6.109), the 'effective multiple scattering' object $\hat{q}_{eff}$ is defined by

$$\exp(-jk\eta)\ \hat{q}_{eff} = s/A_{gen} \ . \tag{6.111}$$

Substituting (6.107), (6.108), and (1.180) into (6.111), the effective multiple scattering object is found to be

$$\hat{q}_{eff} = \hat{q}\left(1 + A/\psi_{gen}\right) \ . \tag{6.112}$$

When monochromatic images $\hat{q}_{eff}(\mathbf{r}|k)$ of the multiple scattering object are combined over a wide bandwidth, the frequency dependence of the A/ψ_{gen} term in (6.112) is usually sufficiently quasi-random that the term does not contribute to the result, and q is obtained. This is the case for elementary shift-and-add as can be seen by substituting (6.112) into (6.101), with $\hat{q}'$ replaced by $\hat{q}_{eff}$.

Some experimental results which suggest that shift-and-add does indeed reduce the artefacts caused by multiple scattering are shown in Figs. 6.13 and 6.14. A target consisting of a metal strip and two wires was constructed using the cassette method described in ¶5.2.3. Scattering measurements were made with the target in three different orientations with respect to the incident beam. The configuration and the orientations of the target are shown in the left column of Fig. 6.13. Direct scattering of the incident wave occurs from the two wires and the edges of the strip (invoking the concepts of the geometrical theory of diffraction mentioned in ¶1.6.2). The specular reflection of the incident wave by the surface of the strip is arranged to miss the measurement aperture given by (5.22) and (5.23). A strong multiple reflection caused by scattering from the smaller wire being reflected by the strip is expected. This multiple reflection is equivalent to a scatterer located at the reflection of the smaller wire in the strip. Since most of the multiple reflected rays are captured by the measurement aperture, the apparent scatterer is expected to have a similar brightness to the smaller wire.

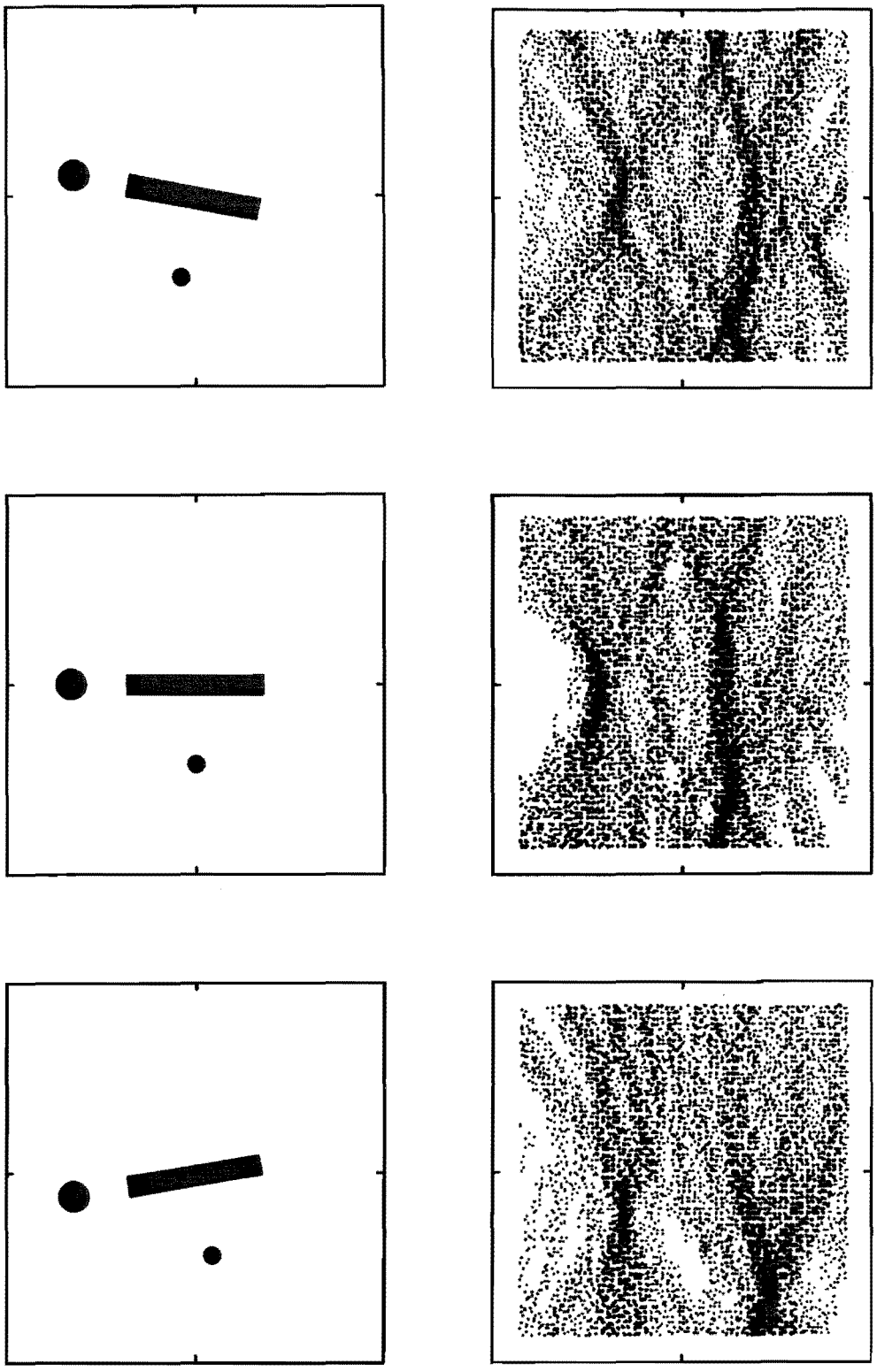


Fig. 6.13 Configurations (left column) and narrow-band images (right column, s_{75}) of a strip-and-two-wire scattering target (5 mm × 0.2 mm brass strip, 0.25 mm and 0.18 mm diameter copper wires). The narrow-band image to the right of each configuration corresponds to it.
Display parameters: $d_o = 0$; $d_\varepsilon = 0$; $d_1 = 1$.

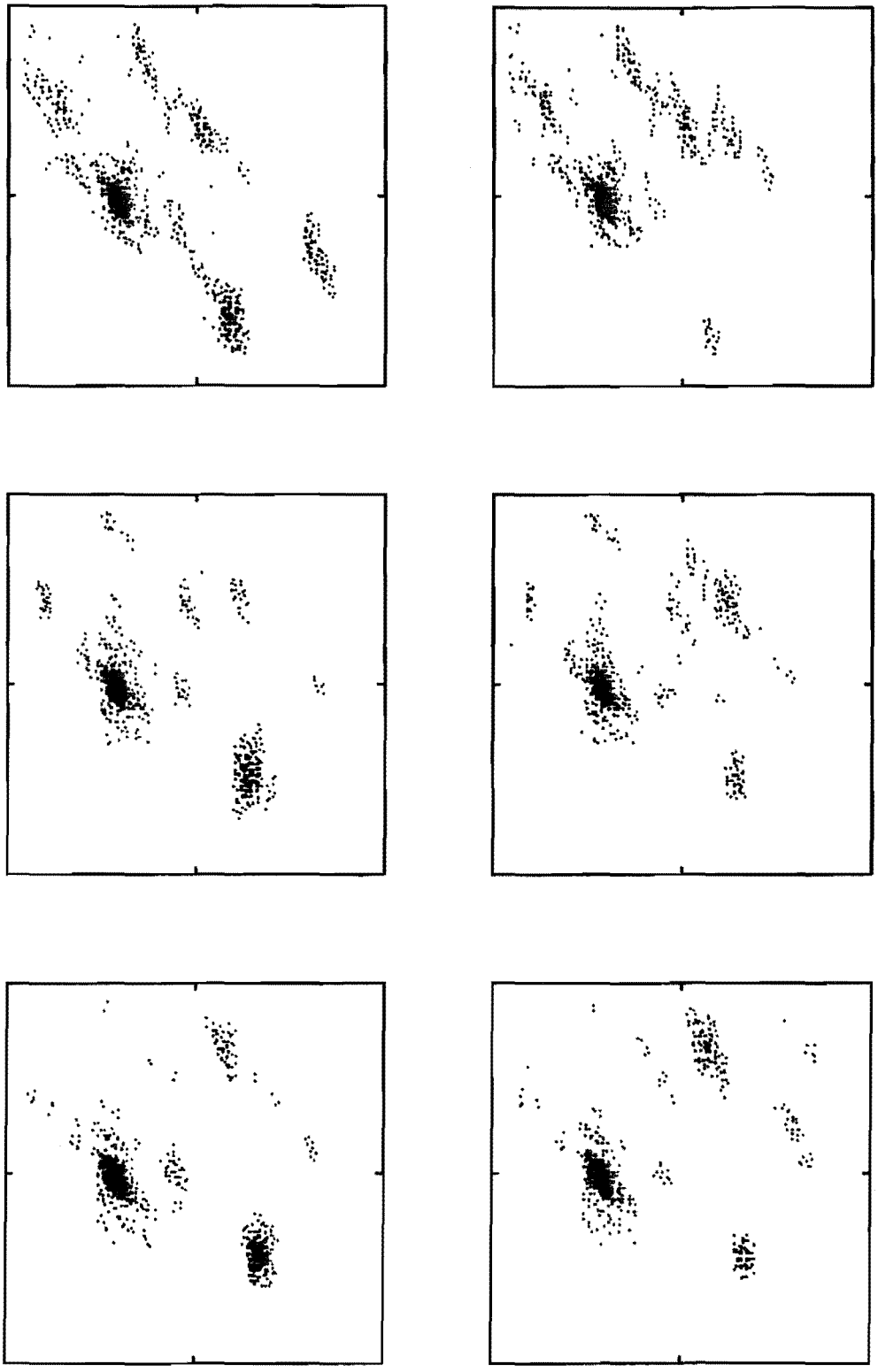


Fig. 6.14 Shift-and-add images of strip-and-two-wire target. The configurations of the target are given by the equivalent rows of Fig. 6.13. Left column: small wire enhanced. Right column: reflection of small wire enhanced.

Display parameters: $d_0 = 0.016$; $d_\varepsilon = 0$; $d_1 = 0.5$.

The image of the larger wire is of similar brightness to that of the smaller wire in the narrow-band images. Even though the images are heavily blurred they are brighter where the scatterers are expected to be located. The edges of the strip are visible in the shift-and-add images. If the small wire has coordinates (x,y) then the phase of the incident wave at the wire is -jkx. Its reflection also has the phase -jkx, but in general it has the coordinates (x',y'). Unless $x = x'$, the phase of the reflection differs from that of the incident wave at (x',y'). Hence shift-and-add is expected to reduce the amplitude of the reflection, provided $x \neq x'$. The case $x = x'$ is illustrated in the middle row of Fig. 6.14. By comparison, in the shift-and-add images for the top and bottom rows the image of the reflection, if not totally removed, is at least spread out. In some of the shift-and-add images ghosts of the edges of the strip and the reflection can be seen. This indicates that the image of the smaller wire is brighter than that of the larger wire in some narrow-band images. This true of the narrow-band images displayed in Fig. 6.13. However, by inspecting the shift-and-add image it is clear that the larger wire is acting as the reference. Thus in most of the narrow-band images the image of the larger wire is brighter than that of the smaller wire.

The brightness of an arbitrary scatterer can be enhanced by changing the value of the acoustic velocity assumed for the image reconstruction. The acoustic velocity is chosen so that the phase of the scatterer and the larger wire are the same. The brightness of the smaller wire is enhanced in the shift-and-add images in the left column of Fig. 6.14, and the brightness of the reflection is enhanced in the images in the right column. The brightness of the image of the smaller wire is expected to be reduced by this processing, and this is indeed the case as can be seen in the right column of Fig. 6.14. The brightness of the ghosts is not affected because the acoustic velocity is only modified when the brightest point of the narrow-band images is where the larger wire is expected to be.

6.8.2 Refraction

Refraction affects Fourier imaging via the object image q defined in the volume source formulation by (1.123). The object image q depends on frequency if the refractive index is not unity. The q reconstructed by Fourier imaging from narrow-band images obtained at several frequencies has this frequency dependence averaged out. In the process of this averaging the object (e.g. o and n defined by (1.113) and (1.116) respectively) may become blurred.

Bates and Millane (1981) and Bates and Minard (1983) discuss an alternative to the volume source formulation for transforming the wave equation (1.112) into one for a homogeneous medium. The alternative, which may be referred to as the 'ray coordinate' formulation, is presented here because it lends weight to the model (6.91).

A coordinate transformation (¶1.2.2) and an amplitude transformation of the wave function are defined by

$$\bar{\nabla} = n\nabla \tag{6.113}$$

and

$$\bar{\psi} = n^{-1/2}\psi \tag{6.114}$$

where the transformed quantity is denoted by placing a bar over the symbol for the untransformed quantity. The wave equation satisfied by ψ is (1.112) which may be written as

$$(\nabla^2 + n^2k^2)\psi = -\sigma\psi - \tau \tag{6.115}$$

with the aid of (1.116) and (1.117). Note that in this subsection σ is defined by (1.113) and does not represent the seeing disc. The transformed wave equation is

$$(\bar{\nabla}^2 + k^2)\bar{\psi} = -\bar{\sigma}\bar{\psi} - \bar{\tau} \tag{6.116}$$

where $\bar{\sigma}$, $\bar{\tau}$, and an inhomogeneity vector $\bar{\mathbf{n}}$ are defined by

$$\bar{\sigma} = \sigma/n^2 - (\bar{\nabla}\cdot\bar{\mathbf{n}} + \bar{\mathbf{n}}\cdot\bar{\mathbf{n}}/2)/2 \ , \tag{6.117}$$

$$\bar{\tau} = n^{-3/2}\tau \ , \tag{6.118}$$

and

$$\bar{\mathbf{n}} = (\bar{\nabla}n)/n \ . \tag{6.119}$$

In the ray coordinate formulation $\bar{\sigma}$ is not frequency dependent. Since (6.116) and (1.122) have the same form, the theory of Fourier imaging (¶1.7) may be developed in the transformed system. The object image is now $\bar{\sigma}$ rather than q. The incident wave which is planar in the transformed coordinates is generated by the sources $\bar{\tau}$. Since the sources of the plane wave in the untransformed formulation are τ the two are unequal. The incident wave in the ray coordinate system is refracted as it enters the object as can be seen from (6.118). In the ray coordinate formulation of Fourier imaging the data $\bar{\chi}$ is obtained from the scattered wave $\bar{\psi}$ in the Born approximation according to (1.222), which is rewritten as

$$\bar{\chi}(\bar{\mathbf{r}},\omega) = \frac{\bar{\psi}(\bar{\mathbf{r}},\omega)}{\bar{A}_{gen}(\omega)\ G_o(\bar{\mathbf{r}},\omega)} \ . \qquad (6.120)$$

However, because the refractive index is unknown, $\bar{\mathbf{r}}$ cannot be calculated given $\mathbf{r}$. Points in the aperture are therefore at unknown distances from the object in the transformed system. Hence, there is an unknown phase relationship between ψ and $\bar{\psi}$, and hence between χ and $\bar{\chi}$. The model (6.91) describes the isoplanatic component of this. Thus shift-and-add imaging takes into account the isoplanatic component of the effects of refraction.

6.9 SUMMARY

In this chapter a basis for shift-and-add imaging is established by theory, by simulation, and by experiment. The procedure for simulating speckle images of an arbitrary object with an arbitrary seeing disc in one or two dimensions allows shift-and-add to be tested by simulation in a wide variety of circumstances. A simple model of the speckle PSF, justified in ¶6.4.1, shows that a speckle image consists of many randomly located and randomly phased replicas of the object. Ideal shift-and-add combines the brightest one of these from each speckle image. Knowledge of the speckle PSF is required for this, however. Weighted shift-and-add does not require this knowledge, but requires that the object contain an isolated bright reference. In some circumstances, naturally occurring biological structures can provide such a reference. When the reference is not a great deal brighter than the rest of the object, an artefact called 'ghosting' occurs. A model is derived on theoretical grounds to explain this artefact and is verified by simulation. It shows that the level of ghosting is determined by the relative integrated amplitudes of the reference and the rest of the object. The ghosting artefact resembles the autocorrelation of the object for positive objects. The model is derived for weak ghosting only.

Shift-and-add when combined with Fourier imaging is an approximate solution to the remote probing problem. Diffraction-limited images of the object are produced in the presence of isoplanatic distortions, provided these are 'speckle-like' and provided the measurements are not too contaminated with noise. Refraction is shown to produce speckle-like distortions with some isoplanatic component.

Ghosting is the chief image artefact created when ideal shift-and-add is approximated by weighted shift-and-add. In Chapters 7 and 8 several methods for reducing the ghosting artefact for certain classes of objects are presented.

7. THE CORRELATION EXTENSION OF SHIFT-AND-ADD

7.1 INTRODUCTION

The chief artefacts of shift-and-add images are caused by ghosting as explained in ¶6.7. In this chapter a technique is presented which attempts to reduce ghosting. Essentially, the simple model of the structure of a speckle image propounded in ¶6.4 is used to derive a more reliable method for determining the shift.

The method is described in ¶7.2 and its implementation as a GEM command module (¶4.8) is explained in ¶7.3. An iterative procedure incorporating the new method is presented in ¶7.4 and the convergence of the method for an object consisting of two delta functions is analysed in ¶7.6. Results based on simulated speckle images and ultrasonic measurements are presented in ¶7.7.

7.2 METHOD

Ghosting (¶6.7) is the worst artefact of weighted shift-and-add. To explain how it occurs the simple model of a speckle image described in ¶6.4.2 is invoked. In this model the speckle image is a superposition of randomly located and scaled replicas of the object. When the object contains a reference (¶6.4.4) the brightest point of the speckle image usually corresponds to the position of the reference in the brightest replica. However, when the object contains a peak other than the reference, it is possible that two or more replicas might reinforce each other at the location of the peak rather than the reference. If they reinforce each other sufficiently, this location becomes the brightest point of the speckle image. When the speckle image is shifted the peak is shifted to where the reference should be and the reference is shifted elsewhere, thus giving rise to a ghost. When the object contains many peaks besides the reference the situation is compounded, as is shown in ¶6.7.1.

In this chapter and Chapter 8, techniques are described which reduce ghosting. The essence of the techniques is the transformation of the speckle images to forms in which the brightest replica of the object may be more readily distinguished. In the conventional technique of weighted shift-and-add the replicas are identified by the reference. In the more general technique proposed in this chapter, a particular estimate

(denoted $\hat{f}$) of the object allows the location of the brightest replica of the object to be identified more surely. The estimate may be obtained from a priori knowledge of the object or more interestingly, from a practical viewpoint, by an iterative scheme described in ¶7.4.

The method which is investigated in this chapter for identifying the brightest replica uses cross-correlation. A transformed speckle image $\bar{s}_m$ is computed from the cross-correlation (defined by (1.34)) of the estimate with the speckle image s_m, i.e.

$$\bar{s}_m = \hat{f} * s_m \ . \tag{7.1}$$

The position $\mathbf{r}_m$ of the brightest replica is given (ideally) by the brightest point of $\bar{s}_m$, i.e.

$$\mathbf{r}_m = b(\bar{s}_m) \ . \tag{7.2}$$

The speckle images are weighted by the amplitude of the transformed speckle images at the brightest point, i.e.

$$w_m = w^{(n)}\left(b(\bar{s}_m)\right) \tag{7.3}$$

where the weighting function is defined by (6.26) and n is an as yet unspecified integer. Finally, the speckle images are combined as for weighted shift-and-add as defined by (6.27), except that $\mathbf{r}_m$ and w_m are given by (7.2) and (7.3) respectively. The result is called the 'correlation' shift-and-add image f_{cor}, which is described concisely by

$$f_{cor} = \langle w^{(n)}\left(b(\hat{f} * s_m)\right)\, T_{-\mathbf{b}(\hat{f} * s_m)}(s_m)\rangle_m . \tag{7.4}$$

When one wishes to emphasise the dependence of f_{cor} upon $\hat{f}$, it is written as $f_{cor}[\hat{f}]$. The result of substituting (6.17) into (7.1) and invoking (1.41) is

$$\bar{s}_m = h_m \odot \bar{f} + \hat{f} * c_m \tag{7.5}$$

where $\bar{f}$ is defined by

$$\bar{f} = \hat{f} * f \ . \tag{7.6}$$

When the model for the speckle PSF (given by (6.9)) is substituted into (7.5), one obtains

$$\bar{s}_m = \sum_l h_{m,l}\, T_{\mathbf{h}_{m,l}}(\bar{f}) + \hat{f} * c \ . \tag{7.7}$$

Hence, the transformed speckle image consists of many replicas of $\bar{f}$. Further more, these replicas have the same locations, i.e. $\mathbf{h}_{m,1}$, as do the replicas of f in s_m (as shown by (6.18), for instance). By determining the position $\mathbf{r}_m$ from (7.2) rather than (6.28), one is locating a replica of $\bar{f}$ instead of f. Thus, it is reasonable to call $\bar{f}$ the 'effective' object (of correlation shift-and-add). When the estimate is identical to the object, $\bar{f}$ is referred to as the 'ideal' effective object (of correlation shift-and-add). From (7.6) the ideal effective object is seen to be the autocorrelation of f. Thus, if $\hat{f}$ is a good estimate of the object, the effective object resembles the autocorrelation of f. Since the autocorrelation of an object always has a peak at the origin, it is postulated that the replicas of the effective object are easier to identify than those of the object itself.

7.3 IMPLEMENTATION

The correlation shift-and-add algorithm has been implemented as a GEM command module (¶4.8) using the software design methodology expounded in Chapter 3. Essentially, the software specification has two sections. Firstly, the module must implement the speckle simulation procedure described in ¶6.3.1. Secondly, it has to implement the correlation shift-and-add algorithm. The 'requirements' (¶3.4.2) for each section comprise, amongst other things, the equations for each procedure.

The command module is called 'cor>'. The syntax of the commands it supports is indicated by the Help display shown in Fig. 7.1. A hyphen appearing at the end of a line indicates that the syntax is continued on the next line. The syntax is specified by the conventions defined in ¶5.5.3.

The calculations are performed by application routines (¶3.4.2) which in turn call GEM library routines. The library routines perform basic image operations (¶1.2.3) such as computing Fourier transforms, conjugating an image, and multiplying or adding two images together. The organisation of the calculations can best be explained by considering the program as an image processing machine. The machine is similar to a general purpose digital computer except that, instead of its registers holding perhaps a 32-bit binary value, they may contain complex images of an arbitrary size. These images are represented with the aid of the image parameter class defined in ¶4.5.1. The operations that the machine can perform on its 'image' registers include multiplication, addition, and

```
Correlation shift-and-add commands:

        Accept/Examine={All,Accept}     field
        Calculate       Correlation             estimate speckle correlation
        Calculate       Estimate                -
                speckle mask_field&ref estimate_(mat,field,point)_id
        Calculate       SAA                     SAA
        Calculate       SAA/Ensemble=num_sca    -
                object  seeing  spec_field      est_par SAA&ref
        Calculate       SAA/Speckle             -
                speckle                         est_par SAA&ref
                        /Estimate(D)    est_par = [estimate&ref cor_field]
                        /Delta          est_par = est_ref
                        /[No(D)]Zero(&reset)/Extend=ext_x,ext_y
        Calculate       Speckle                 object  seeing  speckle
                image = {^Inverse,^Fourier} matrix
                speckle/Speckle=num/Sy=sy
        Image           image   /Image=image/Field=field/Reference=point
                                /Inverse/Fourier
        Peak/Accept(D)  X       Y       A       ACCEPT
        Peak/All        X_ALL   Y_ALL   A_ALL   ACCEPT_ALL
        Random/Seed=seed(&reset)/Phase=phase_max/Dtype=dtype
        Show            {Accept,Images,Peaks,Random,Run,SAA}/format     put_area
        image   = {Object,Seeing,Speckle,SAA,Estimate,Mask,Correlation}
        put_area        = {     Print(I)        just (left I)   print
                                Area            area
                                Field           field           device}
```

Fig. 7.1 Help display of cor> command module.

Fourier transformation. The structure of the machine is especially designed to implement the software specification mentioned previously. This structure is shown in Figs. 7.2 and 7.3. A rectangle represents an image and the name of the image is shown within the rectangle. The exceptions to this rule are the rectangles marked $\mathbf{r}_m$ and a_m which represent the brightest point $\mathbf{r}$ of the transformed speckle image $\bar{s}_m$ and the amplitude of $\bar{s}_m$ at $\mathbf{r}_m$, respectively. The notation of ¶6.3.1 is used to identify the image registers in Fig. 7.2. In Fig. 7.3 the weighted and shifted speckle image is denoted by s'_m. The user of the machine is able to transfer image data to and from any of the registers outlined with the solid rectangles. Dashed rectangles indicate registers which are 'internal' to the machine. The contents of the registers may be observed, however, by operating the machine with special values in the other registers. For example, the

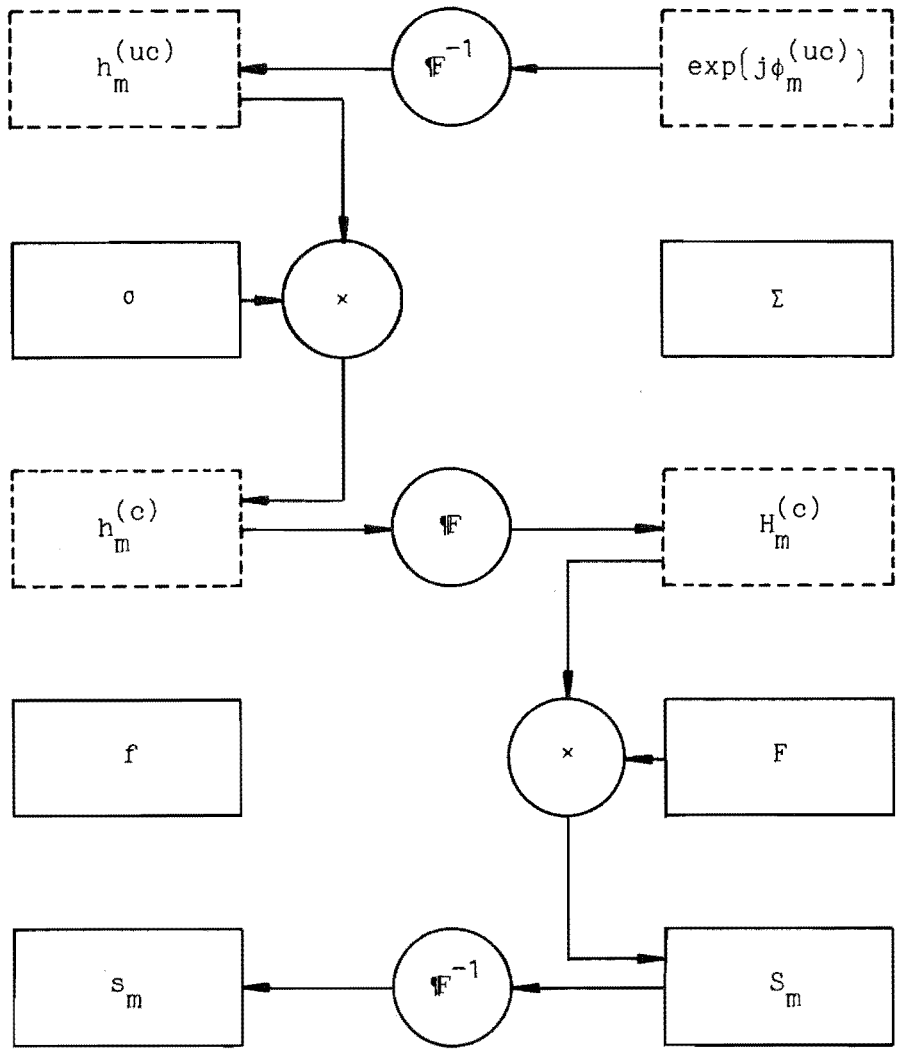


Fig. 7.2 The flow of data for the speckle simulation procedure of ¶6.3.

contents of the internal image register $h_m^{(uc)}$ may be observed by using a uniform seeing disc o and a δ function object. The speckle image register contains $h_m^{(uc)}$ in this case. The operations that the machine performs are indicated by circles. The name of each is shown in the interior of each circle. Arrows indicate which registers supply input data for the operation and where the result is to be placed. As can be seen from the Figs. 7.2 and 7.3, the registers are arranged in two domains. The domain

consisting of the registers in the left hand column is called the 'inverse' domain, while the domain comprising the registers in the right column is called the 'Fourier' domain. These domains are so named because in general the image registers opposite each other in the two domains contain an image and its Fourier transform. The adjective 'inverse' is used for want of a better word. The adjective 'object' would be confusing since there is a register F related to the object in the Fourier domain.

The user controls the machine by issuing the commands listed in Fig. 7.1. The command parameters are related to the image registers. For instance, when loading a GEM image (¶4.5.1) into the $\hat{f}$ register the user specifies the name of the image as the 'estimate' parameter for the appropriate command. In actuality, no transfer takes place because the machine uses a 'pointer' to locate the image data in memory. In addition to specifying the name of the image, the domain of the register which is to be loaded may also be specified. If the calculation specifying the command requires the image to be in the other domain, the machine automatically calculates a forward or inverse Fourier transform as appropriate. The result of the calculation is treated analogously. These features are useful because they relieve the user of the need to specify an image in a particular domain. However, when several commands are to be executed, it is possible to eliminate unnecessary Fourier transformations by specifying the images in the domains required by the machine.

Figs. 7.2 and 7.3 do not indicate the entire capability of the machine. For instance, the f_{cor} register may be zeroed, and the addition of the s'_m image to f_{cor} may be inhibited if the brightest point is not within a specified region of image space. The addition is performed coherently or incoherently depending on whether the data type (¶4.5) of the f_{cor} image is complex or real respectively. A single command can result in a specified number (specified by the /Ensemble qualifier) of speckle images being generated and processed by correlation shift-and-add. The image register pair (s_m, S_m) shown at the bottom of Fig. 7.2 is also shown at the top of Fig. 7.3. Thus the procedures indicated in these two figures may be 'pipelined', i.e. carried out one after the other. Only the weighting function $w^{(0)}$ is implemented by the machine. An elementary shift-and-add image is computed in response to the command 'Calculate SAA/Delta', while the command 'Calculate SAA/Estimate' causes f_{cor} to be calculated.

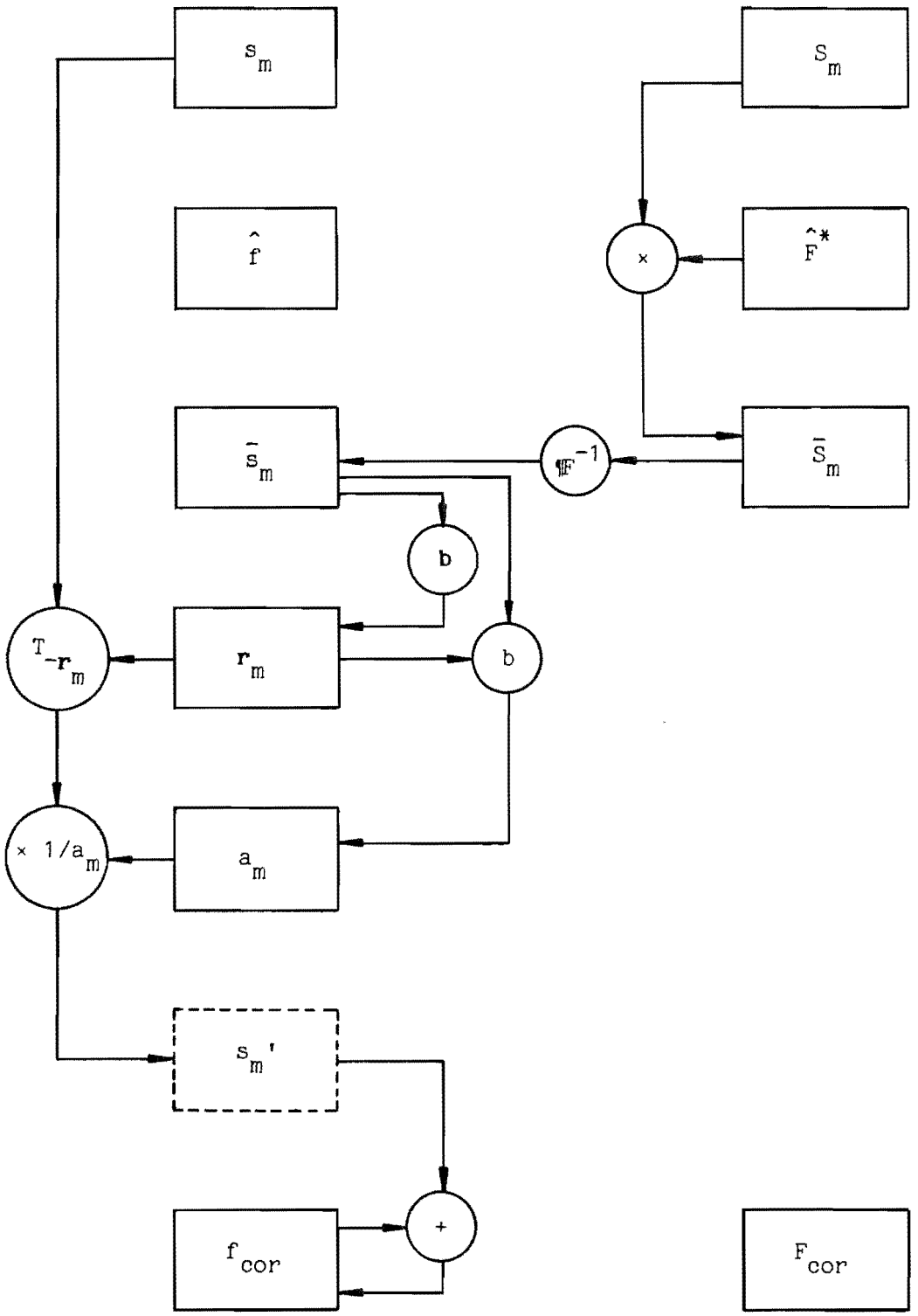


Fig. 7.3 The flow of data for correlation shift-and-add.

7.4 AN ITERATIVE SCHEME

As indicated in ¶7.2 the estimate required by correlation shift-and-add can be constructed iteratively. The first estimate $\hat{f}^{(1)}$ is obtained from ordinary shift-and-add applied to the ensemble $\{s_m\}$. The second estimate $\hat{f}^{(2)}$ is obtained by performing correlation shift-and-add on the ensemble $\{s_m\}$ with $\hat{f}^{(1)}$ as the estimate. In general, the i^{th} estimate $\hat{f}^{(i)}$ is obtained by performing correlation shift-and-add with the estimate $\hat{f}^{(i-1)}$, hence the iterative procedure may be written

$$\hat{f}^{(1)} = f_{wsaa} , \tag{7.8}$$

$$\hat{f}^{(i)} = f_{cor}[\hat{f}^{(i-1)}] \tag{7.9}$$

where the superscript (n) in $f^{(n)}_{wsaa}$ is omitted to prevent confusion with the superscript used to identify the iteration. By noting that

$$f_{wsaa} = f_{cor}[\delta] , \tag{7.10}$$

the iterative scheme may be reformulated in a simpler manner:

$$\hat{f}^{(1)} = \delta , \tag{7.11}$$

$$\hat{f}^{(i)} = f_{cor}[\hat{f}^{(i-1)}] . \tag{7.12}$$

The simulation presented in Fig. 7.4 indicates that the simple iterative scheme specified by (7.11) and (7.12) does not necessarily converge to the object. The particular object f shown in Fig. 7.4(a) is described by

$$f = T_{-\alpha}(\delta) + T_{\alpha}(\delta) \tag{7.13}$$

where

$$\alpha = \hat{x}/32 . \tag{7.14}$$

The magnitude of the ghost peaks in the correlation shift-and-add images $\hat{f}^{(2)}$ and $\hat{f}^{(8)}$ are increased rather than decreased with respect to those in the elementary shift-and-add image f_e. The reason for this behaviour may be explained by considering the effective object for each iteration. First, note from Fig. 7.4 that the elementary shift-and-add image is given approximately by

$$f_e = \tfrac{1}{2}T_{-2\alpha}(\delta) + \delta + \tfrac{1}{2}T_{2\alpha}(\delta) , \tag{7.15}$$

hence the effective object for the second iteration is found, by

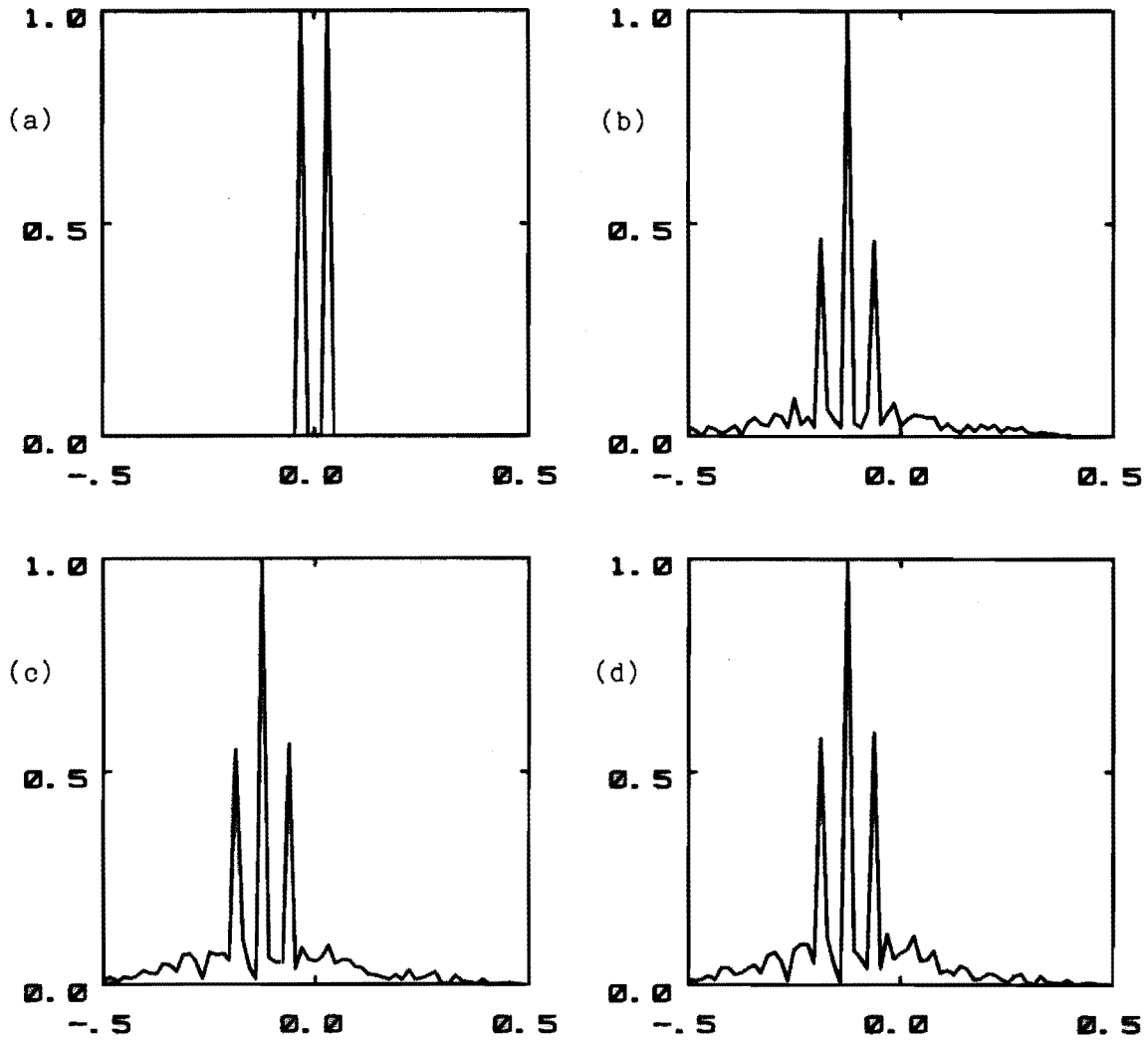


Fig. 7.4 Iterative correlation shift-and-add as defined by (7.5) and (7.6) for a two-point object.
(a) the object f. (b) f_e.
(c) $\hat{f}^{(2)}$. (d) $\hat{f}^{(8)}$.

substituting (7.13) and (7.15) into (7.6), to be

$$\bar{f}^{(2)} = f_e * f$$

$$= \tfrac{1}{2}T_{-3\alpha}(\delta) + \tfrac{3}{2}T_{-\alpha}(\delta) + \tfrac{3}{2}T_{\alpha}(\delta) + \tfrac{1}{2}T_{3\alpha}(\delta) \; . \qquad (7.16)$$

Thus $\bar{f}^{(2)}$ is dominated by two δ functions at $-\alpha$ and α, which have the same relative amplitudes and locations as do the δ functions in the object. It is to be expected therefore that the ghosting in $\hat{f}^{(2)}$ is similar to that in f_e. In fact, it is slightly worse because of the two extra peaks

at -3α and 3α in $\bar{f}^{(2)}$. The ideal effective object (¶7.2) is given by

$$f * f = T_{-2\alpha}(\delta) + 2\delta + T_{2\alpha}(\delta) \ . \tag{7.17}$$

Since the effective object $\bar{f}^{(2)}$ does not have a single dominant peak to act as a reference, as does f * f, the ghosting is not reduced. A solution to this difficulty is derived in ¶7.4.1. It is also obvious from Fig. 7.4 that the average magnitude of the contamination in $\hat{f}^{(8)}$ is greater than in f_e. A method for reducing the contamination in the correlation shift-and-add images is described in ¶7.4.2.

7.4.1 Extent Constraint

The scheme (7.8) and (7.9) must be modified so that the estimate of the object used for each iteration is more faithful. The effective object is then more likely to contain a single dominant peak. The nature of an appropriate modification is suggested by a simple model of the correlation shift-and-add image. This model is an extension of the one proposed in ¶6.7 for the weighted shift-and-add image. Consider the cross-correlation of $\hat{f}$ with (7.4), i.e.

$$\hat{f} * f_{cor} = \langle w^{(n)}\left[b(\bar{s}_m)\right] \, T_{-\mathbf{b}(\bar{s}_m)}(\bar{s}_m) \rangle_m \ . \tag{7.18}$$

The right hand side of (7.18) is the weighted shift-and-add image obtained from the ensemble $\{\bar{s}_m\}$. This is given approximately by the model (6.55). Hence

$$\begin{aligned} \hat{f} * f_{cor} &= p(\bar{f}) * \bar{f} \\ &= \hat{f} * \left[(p(\bar{f}) * f\right] \end{aligned} \tag{7.19}$$

where (7.6) and (1.42) are invoked. Thus

$$f_{cor} = p(\bar{f}) * f \ . \tag{7.20}$$

This is a reasonable formula since the ghosting function p is dependent on the effective object rather than the actual object as in (6.55). The formula (6.89) may be invoked for the ghosting function provided three conditions are met. Firstly, the effective object must contain a bright reference, i.e.

$$\bar{f} = \delta + \bar{g} \ . \tag{7.21}$$

Secondly, the extent of the effective object must be smaller than that of

the seeing disc. Finally, n must be 0 for the weighting function defined by (6.26). When (6.89) is substituted into (7.20), one obtains the formula

$$f_{cor} = \tfrac{1}{2}(1 - \int_{T_{\mathbf{r}}} |\tilde{g}|)f + \tfrac{1}{2}|\tilde{f}| * f \qquad (7.22)$$

for f_{cor}. This formula indicates that f_{cor} contains two components. One is a scaled replica of the object; the other is the ghosting artefact. In the modification to the iterative scheme given by (7.11) and (7.12), an attempt is made to extract the component corresponding to the object. This is done by using an 'extent' constraint which is now described. Note that in general the second term on the right hand side of (7.22) has twice the extent of the first, which is of the same extent as the object. If the first term is larger than the second it should be possible to locate it by examining all regions in image space which have the same extent as the object and by selecting the one which contains the most energy. This region is assumed to correspond to f and is shifted to the centre of image space. The image is set to zero outside the region assumed to be occupied by the object. A significant proportion of the ghosting artefact is expected to be removed by this procedure. The procedure is now described mathematically. Let $\tilde{f}$ denote f_{wsaa} or f_{cor}, and suppose the object to be enclosed by a box $e(\mathbf{r})$. In two dimensions, the box is calculated from

$$e(\mathbf{r}|e_x,e_y,a_x,a_y) = \text{rect}(x/e_x)\,\text{rect}(y/e_y) \circledcirc \text{rect}(x/a_x)\,\text{rect}(y/a_y) \; . \qquad (7.23)$$

This formula corresponds to an 'apodised' box, which has a flat 'top' of dimensions $e_x - a_x$ by $e_y - a_y$ and whose edges fall to zero in a distance a_x from the edge of the flat top in the x direction and a_y in the y direction. The amount c of $\tilde{f}$, contained by the box when it is centred at $\mathbf{r}$, is given by

$$\begin{aligned} c(\mathbf{r}) &= \int_{T_{\mathbf{r}}} e(\mathbf{r}' - \mathbf{r})\; |f(\tilde{\mathbf{r}}')|\; d\mathbf{r}' \\ &= (e * |\tilde{f}|)(\mathbf{r}) \end{aligned} \qquad (7.24)$$

where (1.34) is invoked. The component of f in $\tilde{f}$ is expected to occur at $\mathbf{r}_c$, given by

$$\mathbf{r}_c = \mathbf{b}(c) \; . \qquad (7.25)$$

The estimate $\hat{f}$ is obtained from $\tilde{f}$ by, first, shifting $\tilde{f}$ so that the component of f is centred at the origin and, second, multiplying by the box, i.e.

$$\hat{f} = e\, T_{-r_c}(\tilde{f}) \,. \tag{7.26}$$

7.4.2 Thresholding

The extent constraint of ¶7.4.1 was applied to simulated one and two-dimensional shift-and-add images of an object. The object contained several peaks, one of which was bright enough to act as a reference. The images were displayed on a video monitor.

In each image the object could be identified along with several ghost peaks. In one dimension the extent constraint located the region of the image containing the object. In two dimensions, however, the located region was near the centre of image space, irrespective of the position of the object. Thus in two dimensions the extent constraint, described by (7.23) to (7.26), does not function as intended. The reason for this is now explained. Consider the cross-correlation of the box with the magnitude of the ideal shift-and-add image as given by (6.22),

$$e * |f_{id}| = e * |f| + e * |c_{id}| \,. \tag{7.27}$$

In (7.27) c_{id} is the ideal contamination defined by (6.23). Suppose that in a 64 × 64 sampled representation (¶1.2.5) a single point object has a single nonzero pixel which is of unit amplitude. Further, suppose that the box is 32 × 32 pixels and is not apodised. If $|c_{id}|$ has an average of 0.001 over the entire image the second term on the right hand side of (7.27) is larger than the first. For the shift-and-add images presented in this thesis, which are obtained by averaging 128 speckle images, $|c_{id}|$ is approximately 0.01. Hence, the contamination level c_{id} swamps the cross-correlation. In one dimension, a contamination level 0.03 can be tolerated in the circumstances described above, so the difficulty does not arise. A brute force method for overcoming this difficulty would be to increase the number of speckle images which are averaged. A better method, however, is to apply the threshold operator, defined by (6.39), to $\tilde{f}$ before the cross-correlation is computed. This removes the contamination. Faint detail in $\tilde{f}$ is also removed but this is not important since, if the detail is faint, it does not affect the estimate greatly. Furthermore, the shift determined from the transformed speckle images is unlikely to be affected at all.

In summary, the iterative scheme proposed is:

$$\hat{f}^{(0)} = \delta \,, \tag{7.28}$$

$$\tilde{f}^{(i)} = f_{cor}[\hat{f}^{(i-1)}] , \tag{7.29}$$

$$\tilde{f}_t^{(i)} = thresh(\tilde{f}^{(i)}, \varepsilon_t) , \tag{7.30}$$

$$c^{(i)} = e(\cdot|e_x, e_y, a_x, a_y) * |\tilde{f}_t^{(i)}| , \tag{7.31}$$

$$\hat{f}^{(i)} = e(\cdot|e_x, e_y, a_x, a_y) \; T_{-b[c^{(i)}]}(\tilde{f}_t^{(i)}) . \tag{7.32}$$

The effective object for the i^{th} iteration is given by

$$\bar{f}^{(i)} = \hat{f}^{(i-1)} * f . \tag{7.33}$$

Elementary shift-and-add is used to compute $\tilde{f}^{(1)}$ for all of the simulation results presented in this chapter. To emphasise this, $\tilde{f}^{(1)}$ is also written as $\tilde{f}_e^{(1)}$. In a similar manner $\tilde{f}^{(i)}$ for $i \geq 2$ is written as $\tilde{f}_{cor}^{(i)}$. The limit of one of the sequences of images generated by the iterative scheme (if it exists) is denoted by replacing the index i with the symbol ∞ in the denotation for the member of the sequence. For example, the limit of the sequence $\{\tilde{f}^{(i)}\}$ is denoted by $\tilde{f}^{(\infty)}$.

It is convenient to define an 'extent' image operator E, which represents the application of the extent constraint, by

$$\hat{f}^{(i)} = E\,[\tilde{f}^{(i)}, \varepsilon_t, e] . \tag{7.34}$$

The formula for E is given by (7.30) through (7.32).

For the computations presented in this thesis, most of the additional image processing required to implement the iterative scheme is performed by the general purpose GEM command modules image> and mop>, alluded to in ¶4.8. The cross-correlation in (7.31), however, is performed by the 'Calculate Correlation' command of the cor> module described in ¶7.3. The entire sequence of GEM commands required to implement the iterative scheme is placed in a GEM command file (¶4.7). No additional FORTRAN programming is required.

7.5 ANALYSIS FOR A TWO-POINT OBJECT

In this section the result of applying the iterative scheme (7.28) through (7.32) to a two-point object is examined theoretically. The object f is

$$f = \delta + \alpha T_{\alpha}(\delta) , \quad 0 \leq \alpha \leq 1 \tag{7.35}$$

and the heuristic model (7.22) is assumed. The estimate $\hat{f}^{(i)}$ is assumed to be given by

$$\hat{f}^{(i)} = \delta + \beta^{(i)} T_{\alpha}(\delta) , \quad 0 \leq \beta^{(i)} \leq 1 . \tag{7.36}$$

Note that (7.36) is certainly true for i = 0 because of (7.28). The value $\beta^{(i)}$ is found and the assumption (7.36) is proved correct by invoking the principle of mathematical induction. The estimate $\hat{f}^{(i+1)}$ is now calculated.

The effective object $\bar{f}^{(i+1)}$ is given by (7.33). For the object given by (7.35) and the estimate given by (7.36) it is found to be

$$\bar{f} = \beta T_{-\alpha}(\delta) + (1 + \alpha\beta)\delta + \alpha T_{\alpha}(\delta) \tag{7.37}$$

where $\beta^{(i)}$ is written as β for brevity. To invoke the model (7.22), the unshifted δ function in (7.37) is assumed to be the reference. Also, the effective object is normalised so that the reference has unit strength. The ghosting artefact in (7.22) is given by

$$|f| * f = \alpha T_{-\alpha}(\delta) + (1 + \alpha\beta + \alpha^2)\delta + (\beta + \alpha + \alpha^2\beta) T_{\alpha}(\delta) + \alpha\beta T_{2\alpha}(\delta) \tag{7.38}$$

for the unnormalised reference. Substituting (7.38) and (7.37) into (7.22) with the normalised effective object given by $\bar{f}/(1 + \alpha\beta)$, one finds that

$$f_{cor} = \frac{1}{2(1 + \alpha\beta)}\left[\alpha T_{-\alpha}(\delta) + \left[2 - \alpha + \alpha^2 + \beta(2\alpha - 1)\right]\delta \right.$$
$$\left. + \left[2\alpha - \alpha^2 + \beta(2\alpha^2 - \alpha + 1)\right] T_{\alpha}(\delta) + \alpha\beta T_{2\alpha}(\delta)\right] . \tag{7.39}$$

Application of the extent constraint selects the δ function at **0** and at $\boldsymbol{\alpha}$, hence $\hat{f}^{(i+1)}$ is of the form given by (7.36). When $\hat{f}^{(i+1)}$ is normalised, $\beta^{(i+1)}$ is given by

$$\beta = \frac{2\alpha - \alpha^2 + \beta(2\alpha^2 - \alpha + 1)}{2 - \alpha + \alpha^2 + \beta(2\alpha - 1)} . \tag{7.40}$$

The limit $\beta^{(\infty)}$ of $\beta^{(i)}$ as $i \to \infty$ is found by requiring that

$$\beta^{(\infty)} = \beta^{(i+1)} = \beta^{(i)} = \beta . \tag{7.41}$$

When (7.40) is substituted into (7.41) the result

$$\beta^{(\infty)} = \begin{cases} 1 & \alpha \neq 0 \quad (7.42) \\ \beta^{(0)} & \alpha = 0 \quad (7.43) \end{cases}$$

is obtained. The implications of this result are discussed below. To investigate the convergence of $\beta^{(i+1)}$ to the limit given by (7.42), the error $e^{(i)}$ in the i^{th} iteration is defined by the difference between the

limit and $\beta^{(i)}$, i.e.

$$e^{(i)} = 1 - \beta^{(i)} . \tag{7.44}$$

When (7.44) is substituted for i and for i+1 into (7.40), the result is

$$e^{(i+1)} = \frac{(2\alpha^2 - 3\alpha + 2)\, e^{(i)}}{1 + \alpha + \alpha^2 - e^{(i)}(2\alpha - 1)} . \tag{7.45}$$

For example, if $\alpha = \frac{1}{2}$ then

$$e^{(i+1)} = 4/7\ e^{(i)} . \tag{7.46}$$

Hence, the convergence is like that of a geometric progression.

Consider the case $\alpha \neq 0$. The results (7.36) and (7.42) show that in the limit $\hat{f}^{(i)}$ has the same form as f but that the amplitude of $\beta^{(\infty)}$ is incorrect. This result must not be taken as an exact indication of what occurs in practice for two reasons. Firstly, the model (7.22) assumes that the effective object is dominated by a δ function. However, this is certainly not the case when $\beta^{(i)} \to 1$ since (7.37) contains a term of relative amplitude $\beta/(1 + \alpha\beta)$. Therefore, one can expect that $\beta^{(i)}$ converges to a value less than unity but greater than α. Secondly, the model (7.22) assumes that the seeing disc is wide compared to the object. When it is of an extent comparable to the object, ghosting can be expected to be reduced, causing $\beta^{(\infty)}$ to be smaller than indicated by (7.42). These speculations are verified by the simulations presented in ¶7.6.

When $\alpha = 0$, (7.43) suggests that the estimate $\hat{f}^{(i)}$ remains constant. This is certainly correct when $\hat{f}^{(0)}$ is given by (7.28), since $\beta^{(0)} = 0$ then. However, if the initial estimate has $\beta^{(i)} = 1$, say, then (7.43) indicates that $\beta^{(\infty)} = 1$ also. Since the model is not necessarily valid for $\beta^{(i)} = 1$, one has reason to expect that $\beta^{(i)}$ is reduced after each iteration. This is indeed the case, as the simulation results in ¶7.6 confirm.

7.6 RESULTS

7.6.1 Object Specification

Objects for the simulations presented in this thesis are generated by a GEM command module (¶4.8) called 'synthesis>'. The user of the command module specifies the object image using a symbolic representation (¶1.2.4). The image f is represented by the formula

$$f(x,y) = \sum_{k=1}^{K} c_k(x,y) \tag{7.47}$$

where c_k is called a 'component'. A component is given by

$$c_k(x,y) = A_k \; X_k\left(\frac{x - x_k}{\Delta x_k}\right) \; Y_k\left(\frac{y - y_k}{\Delta y_k}\right) \tag{7.48}$$

where A_k is called the 'amplitude', X_k and Y_k are called the x and y 'profile' functions respectively, x_k and y_k are real numbers called the x and y profile 'centres', and Δx_k and Δy_k are real numbers called the x and y profile 'widths'. A profile function is either a Gaussian, rectangle, triangle, sinc, or δ function. These functions are defined in the glossary. By invoking the definitions of the translation operator (1.31) and the scaling operator (1.32), (7.48) may be written as

$$f(x,y) = \sum_{k=1}^{K} A_k \; T\left(x_k \hat{\mathbf{x}} + y_k \hat{\mathbf{y}}\right)\left[S_{\Delta x_k, \Delta y_k}\left(X_k(x) \; Y_k(y)\right)\right] . \tag{7.49}$$

```
Synthesis commands are:

      Add             synth   amp     x_prof x_centre x_width -
                                      y_prof y_centre y_width
            (the default for the y profile is the x profile)
      Calculate       synth   field   matrix
      Declare         synth   max_num
      Delete          synth   num
      Release         synth
      Set             synth   num     Amplitude
       "                "       "      {X,Y}    Profile profile
       "                "       "        "      Centre  centre
       "                "       "        "      Width   width
      Show            synth   prec    cols     title   /print
```

Fig. 7.5 Help display of synthesis> command module.

To describe an image the user specifies A_k, X_k, x_k, Δx_k, Y_k, y_k, Δy_k for k = 1, ..., K. The collection of these values is called a 'synthesis'. It is implemented as a GEM parameter class (¶4.5) which is based upon the matrix parameter class. The Help display for the synthesis> module is shown in Fig. 7.5. Several of these entities may be created by the user with the

Declare command. They are each referred to by a name specified by the user. The Add command increases K by unity and allows the user to specify an additional component. Any of the attributes of an existing component may be altered by the Set command. A component c_k may be removed from the sum in (7.47) using the Delete command. In Fig. 7.5, k is denoted by 'num'. A synthesis is displayed as a table by the Show command.

The symbolic representation of (7.47) and (7.48) is converted to a sampled one by the Calculate command. This command has a parameter which specifies a GEM image (¶4.5.1) which is used to store the sampled representation. The formula (7.47) is evaluated for each pixel of the sampled image. The finite precision of real arithmetic on a digital computer is taken into account to ensure that

$$\text{rect}(-\tfrac{1}{2}) = 1 \tag{7.50}$$

and

$$\text{rect}(\tfrac{1}{2}) = 0 \ . \tag{7.51}$$

Also the widths specified for the δ function are ignored. The width is taken as the sampling interval Δ in either the x or y directions as appropriate. The δ function is implemented as a rectangle function with an amplitude of $1/\Delta$ so that (1.124) holds.

7.6.2 Simulation Strategy

Simulation is ideally suited to predicting the behaviour of a system in a known initial state when it is disturbed or altered in some manner. When the initial state is unknown, simulation is of lesser use. The latter situation is similar to that encountered when an image reconstruction procedure is investigated by simulation. One cannot determine how the procedure behaves for all objects, only for a finite number of them. For procedures in which the artefacts in the reconstructed image are independent of the object, a simulation of a 'representative' object is often sufficient. However, when the artefacts are highly dependent on the object, as with shift-and-add, the situation is particularly difficult. One method of proceeding is to perform simulations for objects chosen in an ad hoc manner. The deficiency of this method is that special cases may be chosen for which the reconstruction procedure behaves atypically. It is therefore difficult to draw general conclusions from such simulations. In this thesis, simulations are used to support general hypotheses. To do this the simulations are performed on an ensemble of objects. This technique is employed in ¶6.7.4 where two

ensembles of objects of the general form

$$f = \delta + \alpha T_{\boldsymbol{\alpha}}(\delta) \tag{7.52}$$

are simulated. In one simulation, the results of which are shown in Fig. 6.12(a), α assumes 21 values between 0 and 1. Fig. 6.12(b) displays results for a simulation in which α is held fixed and $\boldsymbol{\alpha}$ is given 32 values.

7.6.3 Ensembles of Objects

The ensembles of one-dimensional objects used for the simulation studies reported in this thesis are denoted by DD(α), DR(α), RRE(α), RRS(α), DDD(α,β,γ) and D13(α,β,γ). These objects are defined by:

$$DD(\alpha) = \delta + \alpha T^{\wedge}_{\mathbf{x}/4}(\delta) \ , \tag{7.53}$$

$$DR(\alpha) = \Delta T^{\wedge}_{-\mathbf{x}/8}(\delta) + \alpha T^{\wedge}_{\mathbf{x}/16}\left[S_{1/8}(\text{rect})\right] \ , \tag{7.54}$$

$$RRE(\alpha) = S_{1/8}(\text{rect}) + \alpha T^{\wedge}_{\mathbf{x}/8}\left[S_{1/8}(\text{rect})\right] \ , \tag{7.55}$$

$$RRS(\alpha) = S_{1/16}(\text{rect}) + \alpha T^{\wedge}_{\mathbf{x}/8}\left[S_{1/16}(\text{rect})\right] \ , \tag{7.56}$$

$$DDD(\alpha,\beta,\gamma) = \alpha T^{\wedge}_{-\mathbf{x}/8}(\delta) + \beta\delta + \gamma T^{\wedge}_{\mathbf{x}/8}(\delta) \ , \tag{7.57}$$

$$D13(\alpha,\beta,\gamma) = \alpha T^{\wedge}_{-\mathbf{x}/8}(\delta) + \beta T^{\wedge}_{-\mathbf{x}/16}(\delta) + \gamma T^{\wedge}_{\mathbf{x}/8}(\delta) \tag{7.58}$$

where Δ is the sampling increment in the x direction. For the simulations reported in this thesis,

$$\Delta = 1/64 \ . \tag{7.59}$$

Operator notation is employed for (7.53) through (7.58) rather than algebraic formulas because it is then easier to visualise what the object looks like. For instance, to determine the appearance of an RRE(α) object from

$$RRE(\alpha)(x) = \text{rect}(8x) + \alpha\ \text{rect}(8x - 1), \tag{7.60}$$

some mental arithmetic is required to obtain the width and position of the rectangle functions, whereas from (7.55) one can visualise the object directly as 'two rectangle functions of width $\frac{1}{8}$, one of which is of height α and shifted by $\frac{1}{8}$ units in the x direction'. The notation for the object ensemble is intended to convey the form of the object. The letters D, R, E and S stand for 'δ function', 'rectangle function', 'one eigth' and 'one sixteenth' respectively. Typical members of each ensemble are shown in Fig. 7.6.

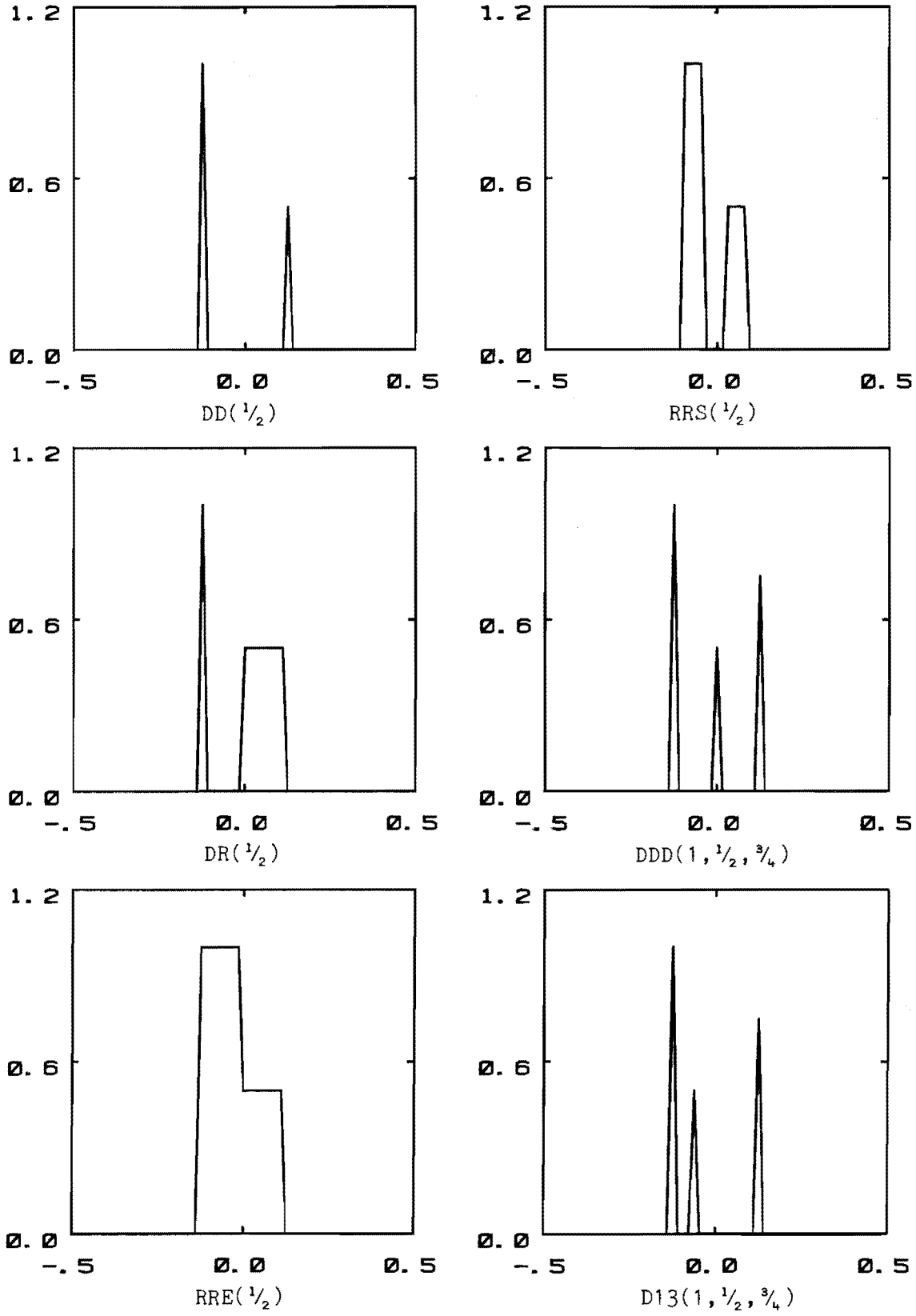


Fig. 7.6 Typical members of each of the object ensembles used for the simulation.

The rationale for the choice (7.53) through (7.58) is now explained. Simple objects are chosen to allow the simulation results to be interpreted as easily as possible. The intention of the simulation is to investigate the efficacy of correlation shift-and-add for reducing the ghosting artefact. Therefore, objects in which the ghosting artefacts arise in a simple manner are chosen. In general, these are objects containing references. DD(α) is the simplest non-trivial such object. DR(α) is a simple object with a reference and an extended background. RRE(α) is an extended object with no reference. RRS(α) is similar to RRE(α) but the two components are isolated rather than coalesced. DDD(α,β,γ) and D13(α,β,γ) are three-point objects. The component spacings in DDD(α,β,γ) are 'redundant', i.e. they are all integer multiples of $\hat{\mathbf{x}}/8$. This is significant because when ghosting occurs, the ghost may appear at the same position as an actual component. In D13(α,β,γ) the component spacings are in a 1:3 ratio and hence only the reference is expected to reinforced by ghosting. These comments are illustrated by Fig. 7.7 which shows the expected shift-and-add images of DDD(1,1,1) and D13(1,1,1). The extent of each of the objects (7.53) through (7.58) is $\frac{1}{4}$, with the exception of RRS(α) for which the extent is 3/16. Hence, a box given by (7.23), with parameters $e_x = 0.4$ and $a_x = 0.1$, has a flat top of length 0.3 which completely includes all of the objects with no truncation due to the apodisation at the edges of the box. Unless specified otherwise, this box is used for all of the simulation results presented in this thesis. The threshold level ε_t (see (7.30)) for the results presented in this chapter is given by

$$\varepsilon_t = 0.1\ b(\bar{f}) \ . \tag{7.61}$$

The ensembles corresponding to each of the formulas (7.53) through (7.58) are generated by varying α, β and γ. For all of the ensembles α, β, γ are real. For the two-component objects DD(α), RRE(α), and RRS(α), α has one of the values in $\{-1, -\frac{1}{2}, \frac{1}{2}, 1\}$, i.e. negative objects are considered. For DR(α), α is 1/16, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$ or 1. For the three-point objects only positive objects are simulated. This so because with three parameters α, β, γ to vary, there are considerably more cases to consider and it was felt that it was better to concentrate on positive objects. If α, β, and γ are chosen from $\{\frac{1}{2}, 1\}$ there are eight possibilities for the triple (α,β,γ). However, some of the objects corresponding to these triples are trivially related. For instance, since

$$\mathrm{DDD}(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) = \tfrac{1}{2}\mathrm{DDD}(1,1,1) \ , \tag{7.62}$$

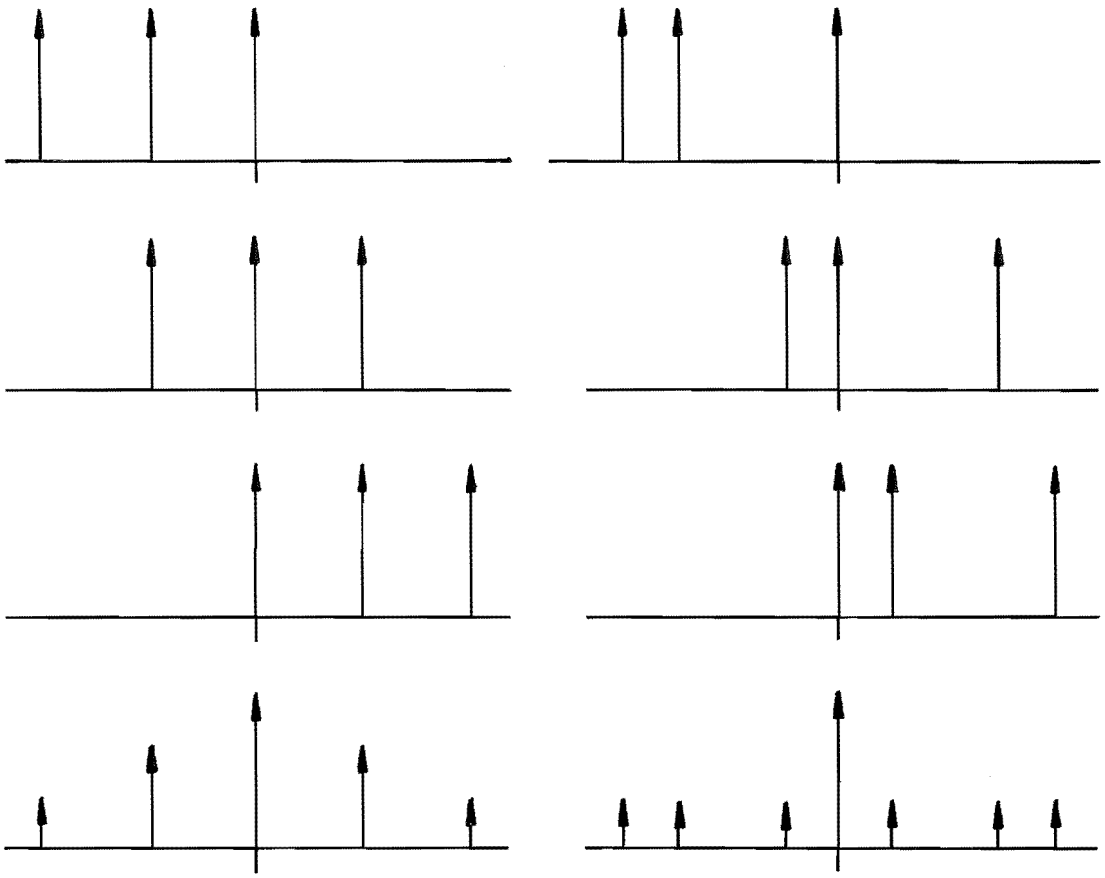


Fig. 7.7 The expected shift-and-add images of DDD(1,1,1) (left column) and D13(1,1,1) (right column). It is assumed that since all of the components are of equal amplitude, that each is chosen as the reference with equal probability. Hence, the shift-and-add image (bottom row) is the sum of the three possible shifted images shown in the top three rows.

the result of simulating DDD($\frac{1}{2},\frac{1}{2},\frac{1}{2}$) can be inferred from the simulation of DDD(1,1,1). For the same reason the D13($\frac{1}{2},\frac{1}{2},\frac{1}{2}$) object need not be simulated. Since

$$\mathrm{DDD}(\tfrac{1}{2},1,1)(\mathbf{r}) = \mathrm{DDD}(1,1,\tfrac{1}{2})(-\mathbf{r}) \tag{7.63}$$

and

$$DDD(\tfrac{1}{2},\tfrac{1}{2},1)(\mathbf{r}) = DDD(1,\tfrac{1}{2},\tfrac{1}{2})(-\mathbf{r}) \ , \tag{7.64}$$

$DDD(\frac{1}{2},1,1)$ and $DDD(\frac{1}{2},\frac{1}{2},1)$ need not be simulated either. Hence, there are five independent possibilities for $DDD(\alpha,\beta,\gamma)$ and seven for $D13(\alpha,\beta,\gamma)$. These are illustrated in Fig. 7.8.

7.6.4 DD(α) Object Ensemble

The result of applying iterative correlation shift-and-add to DD(1) are shown in Figs. 7.9, 7.10, and 7.11. Note that the first iteration corresponds to elementary shift-and-add. The image $\tilde{f}_e^{(1)}$ indicates the chief deficiency of elementary shift-and-add, i.e. its inability to distinguish between points of an object having comparable brightness. On average each point of DD(1) is equally likely to be chosen, but because of the random nature of the speckle images one of the side peaks is always larger than the other. In this instance it is the left hand one. This is detected by the algorithm when the brightest point of $c^{(1)}$ is found. The random low-level contamination in the shift-and-add images is removed by the threshold operation. The result of the next iteration $\tilde{f}_{cor}^{(2)}$ is already a better image of DD(1). The ghost is reduced while the other two peaks are of similar brightness. With further iterations, this trend continues. The effective object converges to the autocorrelation, and therefore has a more clearly defined reference peak. The results for DD(-1) are very similar to those for DD(1).

The convergence of $\hat{f}^{(i)}$ to f can be modelled by (7.40). For $\alpha = 1$,

$$\beta^{(i+1)} = \frac{1 + 2\beta^{(i)}}{2 + \beta^{(i)}} \ . \tag{7.65}$$

The values of $\beta^{(i+1)}$ predicted by (7.65) and the values obtained by simulation for DD(1) and DD(-1) are graphed in Fig. 7.12. The largest difference between theory and simulation is 0.05, so the theory accounts well enough for the observed convergence.

In Fig. 7.13 results are shown for the remaining members of the DD(α) ensemble. In all cases the iterations converge and the ghosting is reduced. The theoretical prediction for the values of $\beta^{(i)}$ are obtained from (7.40) for the case $\alpha = \frac{1}{2}$. They may be calculated from

$$\beta^{(0)} = 0 \tag{7.66}$$

and

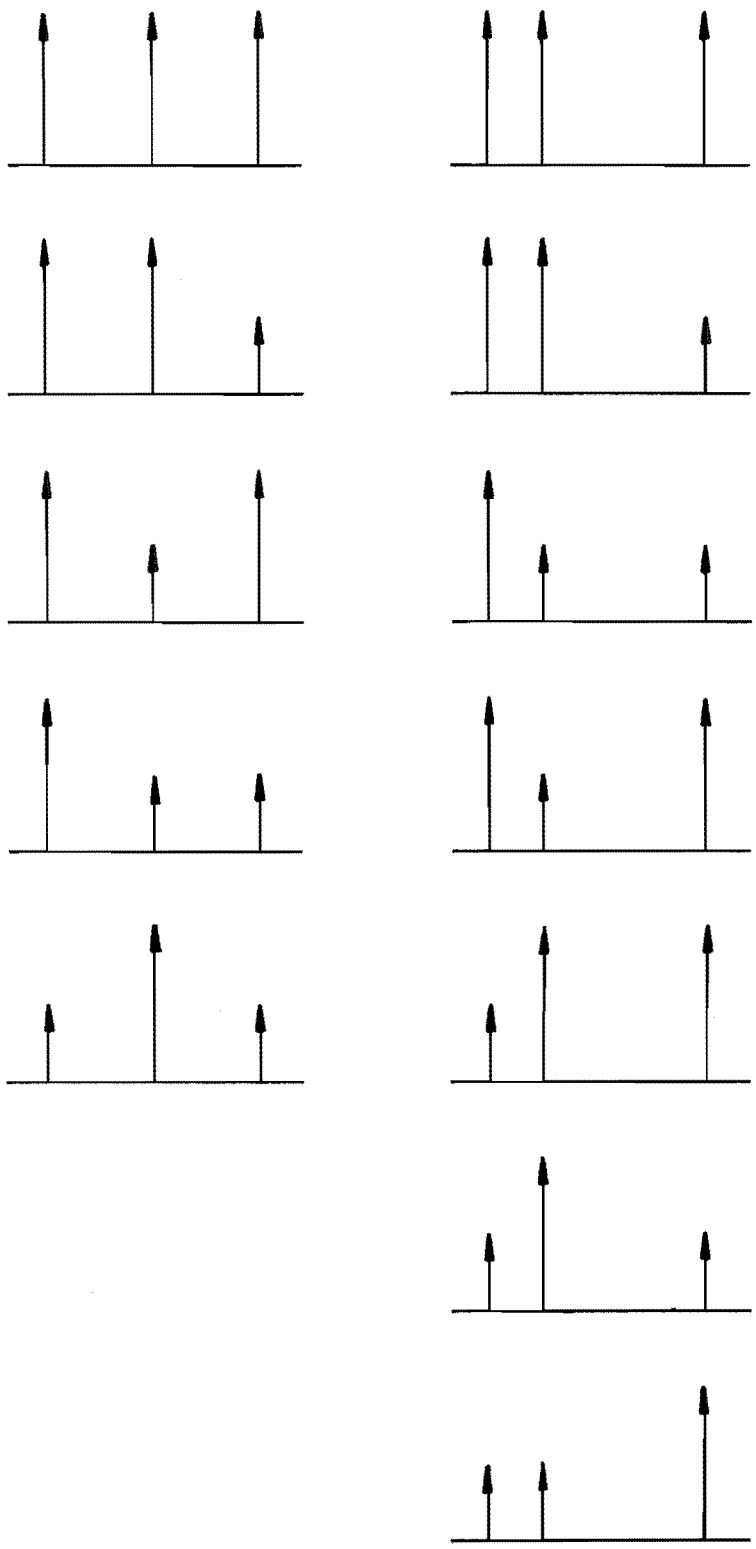


Fig. 7.8 The three-point objects used in the simulation.
Left column: DDD(α,β,γ).
Right column: D13(α,β,γ).

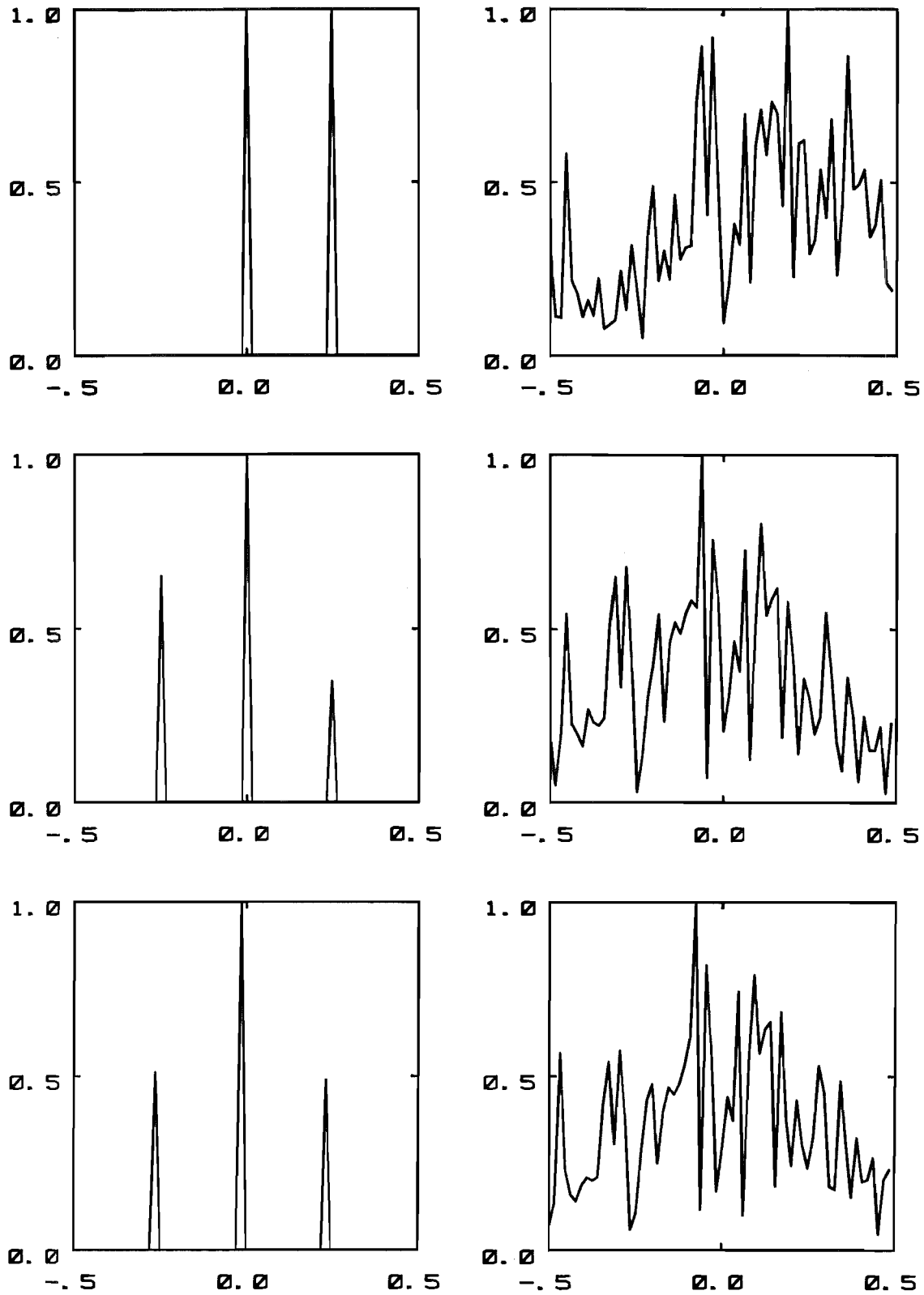


Fig. 7.9 Magnitude of the effective objects $\bar{f}^{(i)}$ (left column) and typical transformed speckle images $\bar{s}_i^{(i)}$ (right column) for the object DD(1).
Top row: i = 1. Middle row: i = 2.
Bottom row: i = 8.

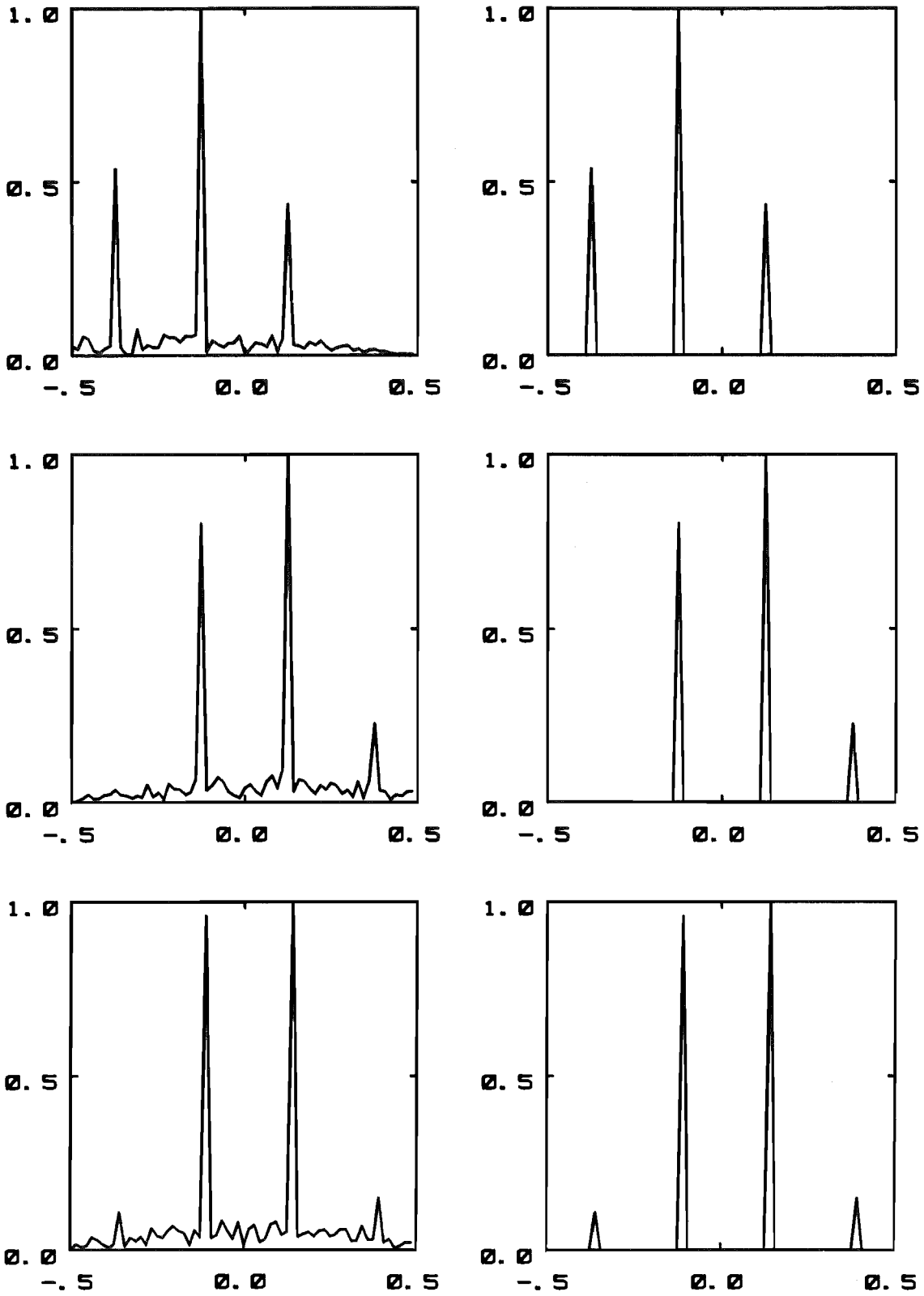


Fig. 7.10 Correlation shift-and-add images $\tilde{f}^{(i)}$ (left column) and thresholded versions $\tilde{f}_t^{(i)}$ (right column) of these for the object DD(1). The rows correspond to the iterations specified in the caption for Fig. 7.9.

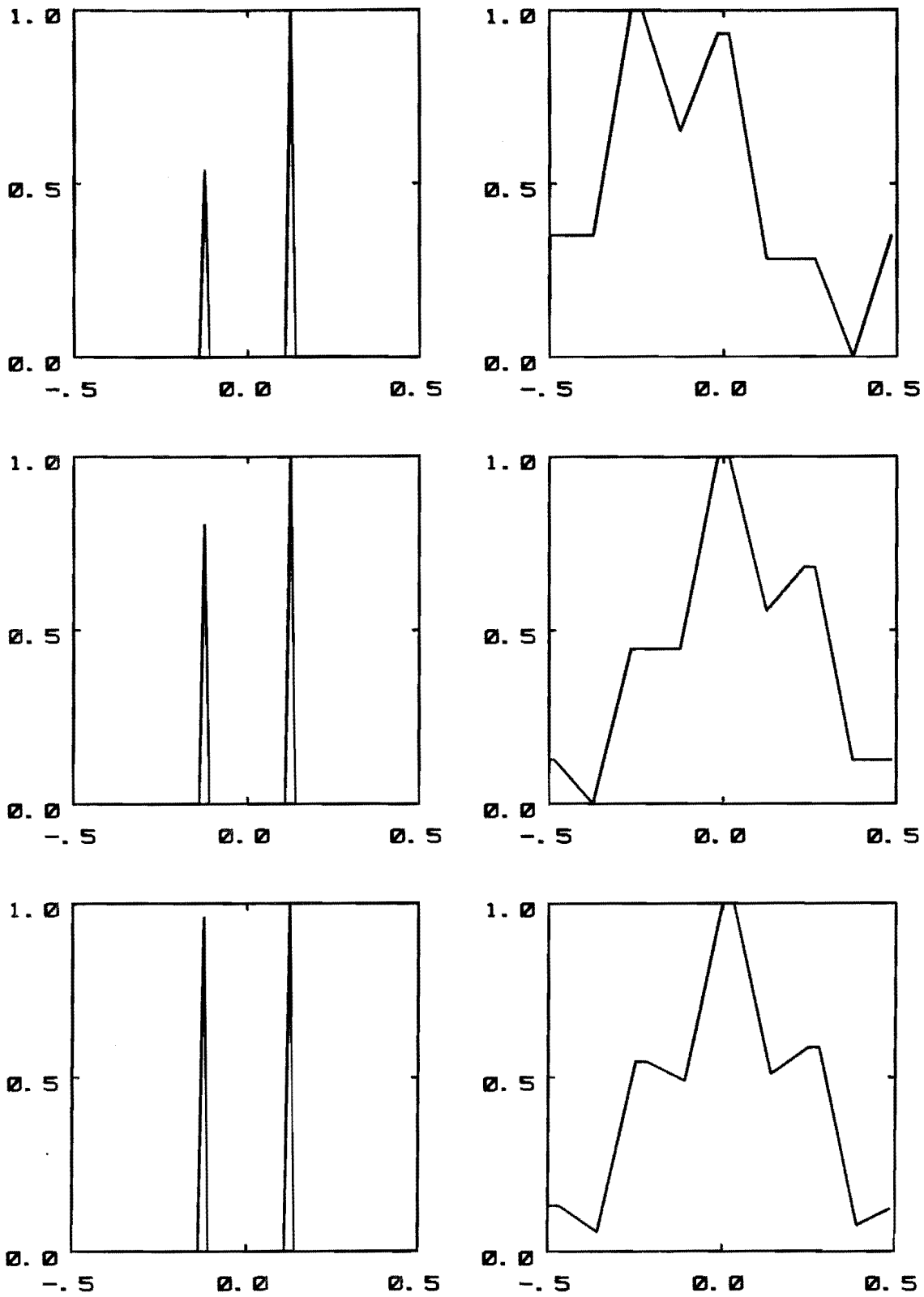


Fig. 7.11 Amplitude distributions $c^{(i)}$ (left column) of $\tilde{f}_t^{(i)}$ and the estimates $\hat{f}_t^{(i)}$ (right column) for the object DD(1). The rows correspond to the iterations specified in the caption for Fig. 7.9.

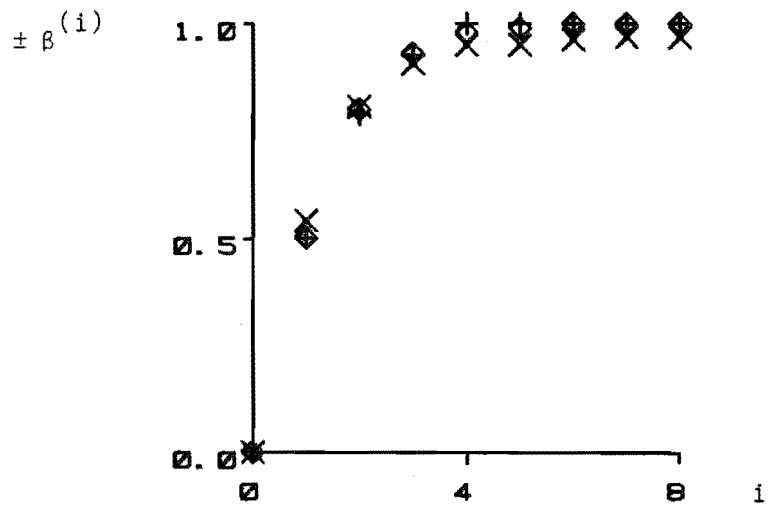


Fig. 7.12 Convergence of the estimate to the object for DD(1) and DD(-1).
Diamonds: prediction of model (7.65).
Crosses: $\beta^{(i)}$ for DD(1) simulation.
Plus signs: $-\beta^{(i)}$ for DD(-1) simulation.

$$\beta^{(i+1)} = (3 + 4\beta^{(i)})/7 . \qquad (7.67)$$

The theoretical predictions and the values of $\beta^{(i)}$ obtained from the simulation are compared for the object DD($\frac{1}{2}$) in Fig. 7.14. For these simulations, it is clear from Fig. 7.14 that $\beta^{(i)}$ does not converge to the actual value of $\alpha = \frac{1}{2}$ nor does it converge to the limit of unity predicted by the model, but to a value in between. This value is 0.69 for DD($\frac{1}{2}$). The case for DD($-\frac{1}{2}$) is similar, with the limit $\beta^{(\infty)} = -0.65$. This confirms the speculations made in ¶7.5.

Consider the situation where the object is a δ function and the estimate is DD(1). In this case the a priori information about the object is incorrect in that it indicates that the object has a peak at $\hat{x}/4$ when in fact it does not. This situation is discussed in ¶7.6. There it is speculated that the extra peak in the estimate is reduced to a constant

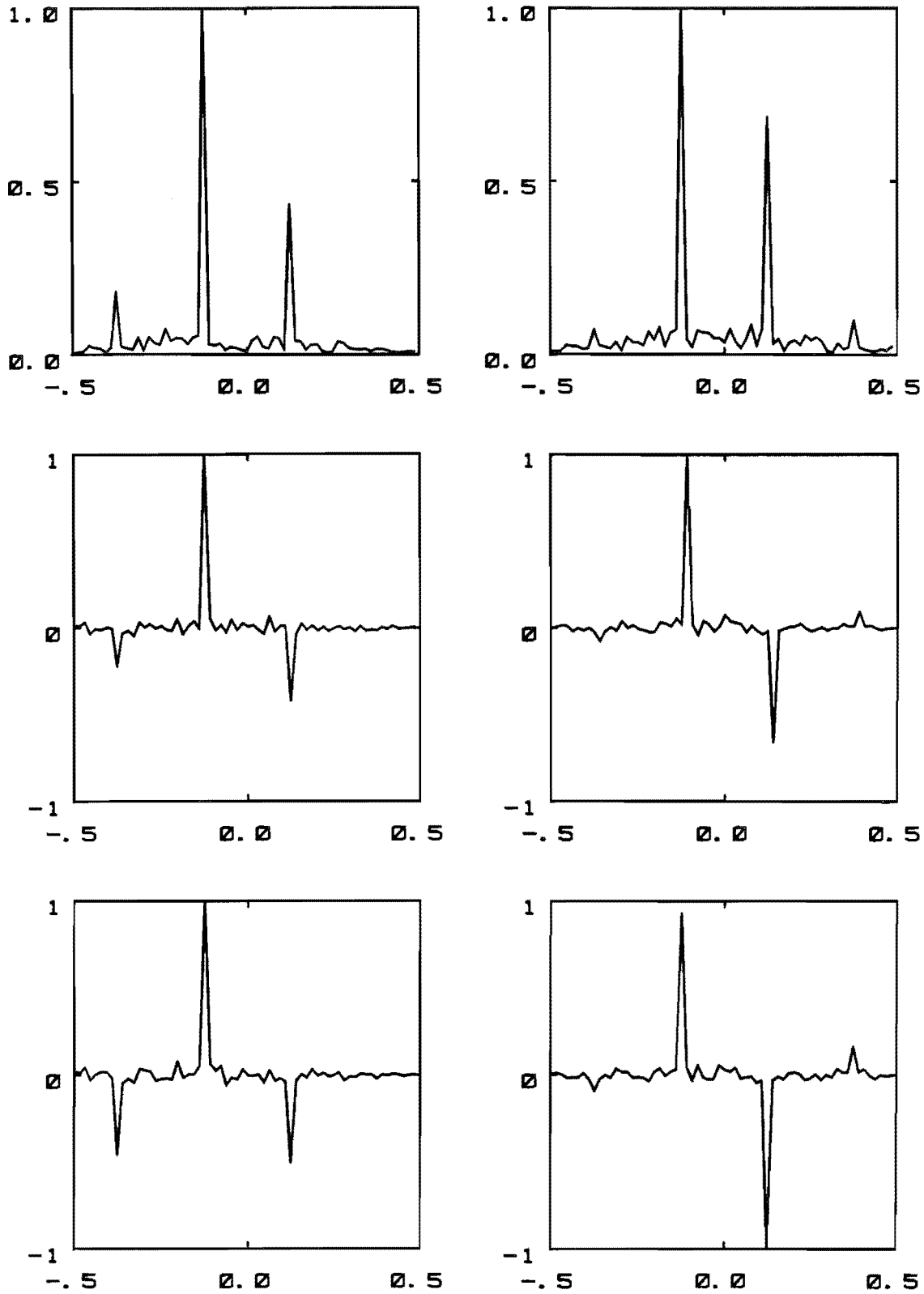


Fig. 7.13 Elementary shift-and-add and correlation shift-and-add images.
Left column: $\tilde{f}_e^{(1)}$. Right column: $\tilde{f}_{cor}^{(8)}$.
Top row: DD($\frac{1}{2}$). Image magnitude displayed.
Middle: DD($-\frac{1}{2}$). Real part of image displayed.
Bottom row: DD(-1). Real part of image displayed.

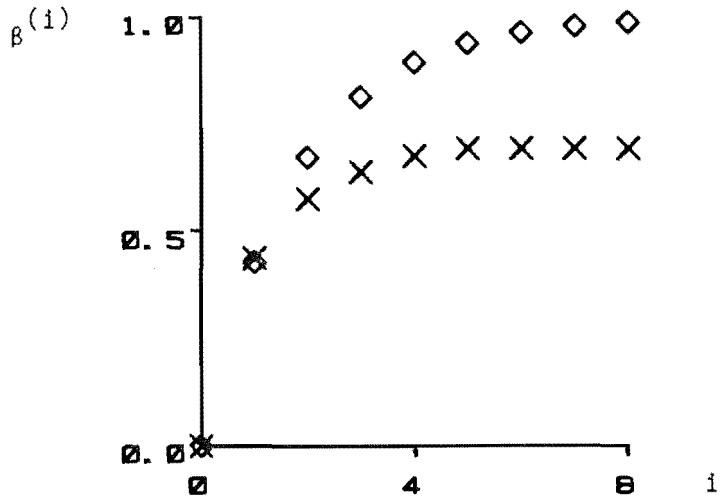


Fig. 7.14 Convergence of the estimate to the object for DD(½).
Diamonds: prediction of model (7.67).
Crosses: $\beta^{(i)}$ for DD(½) simulation.

nonzero amplitude. This is confirmed by the values of $\beta^{(i)}$ obtained by simulation which are graphed in Fig. 7.15. The value of $\beta^{(i)}$ drops to 0.125 for the fourth iteration and then remains approximately constant. Thus, incorrect a priori knowledge can introduce ghosts. Consider the situation in which the threshold level is zero. All of the background contamination in the image $\tilde{f}^{(i)}$ is transferred to the estimate $\hat{f}^{(i)}$. The above result suggests that the background contamination is increased to a level of approximately 0.125. An example where a peak belonging to the contamination is enhanced in this way is discussed in ¶7.6.6.

7.6.5 DR(α) Object Ensemble

For an object of the form

$$f = r + g \tag{7.68}$$

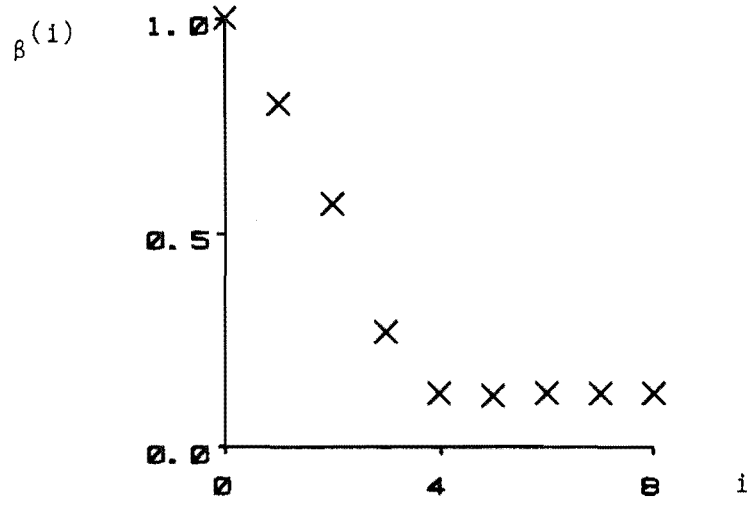


Fig. 7.15 The effect of incorrect a priori knowledge. The object is a δ function but the initial estimate is DD(1).

where r is an isolated reference, the model (7.22) indicates that the 'level of ghosting' μ_ℓ is given by the ratio

$$\mu_\ell = \frac{\int_{T_{\mathbf{r}}} |g| \, d\mathbf{r}}{\int_{T_{\mathbf{r}}} |r| \, d\mathbf{r}} \tag{7.69}$$

which is interpreted in the following way. If $\mu_\ell << 1$, the ghosting artefact is negligible; if $\mu_\ell \approx 1$, the ghosting artefact is significant; and if $\mu >> 1$ the image is dominated by ghosting.

Now consider DR(α) as given by (7.54). It may be expressed in the manner of (7.68) by defining

$$r = \Delta T_{-\mathbf{x}/8}(\delta) \tag{7.70}$$

and

$$g = \alpha T_{-\mathbf{x}/16}\left(S_{1/8}(\text{rect})\right) . \tag{7.71}$$

The level of ghosting is obtained by substituting (7.70) and (7.71) into (7.69) and invoking (7.59). The result is

$$\mu_{\ell} = 8\alpha . \tag{7.72}$$

This result may be obtained less formally by adding up the pixel amplitudes in r and g. The reference is one pixel of unit amplitude and the function g has $(1/8)/\Delta = 8$ pixels of amplitude α. From (7.72), the shift-and-add images of DR(1/16) should have minimal ghosting, those for DR($1/8$) are expected to have significant ghosting, and those for $\alpha = 1/4$, $1/2$, and 1 should be considerably degraded.

Fig. 7.16 presents results which confirm these predictions. In the elementary shift-and-add image $\tilde{f}_e^{(1)}$ of DR(1/16), the component g is just visible above the background. The threshold level, given by (7.61), exceeds the amplitude of the image of g in $\tilde{f}_e^{(1)}$, hence the estimate $\hat{f}^{(1)}$ is a δ function. Thus, correlation shift-and-add produces an identical result $\tilde{f}_{cor}^{(8)}$ to elementary shift-and-add. For DR($1/8$), the ghost of g in $\tilde{f}_e^{(1)}$ is clearly visible to the left of the central peak. In the $\tilde{f}_{cor}^{(8)}$ image, the ghost is removed (or is at last appreciably smaller than the background). However, an additional ghost has appeared to the right of the image of g. Also, the image of g has an average amplitude of 0.3 which is considerably higher than the actual average amplitude of g which is 0.125.

It is interesting to note, however, that although the amplitude of the correlation shift-and-add image is incorrect by more than a factor of two, the location of g relative to the reference is correct and is in fact all the more clearly indicated by the increased amplitude. Furthermore, although the image of g is increased in amplitude, the amplitude of the ghosts is not. This indicates that actual components of the object may be located more readily in the correlation shift-and-add image than in the elementary shift-and-add image.

These remarks are supported by the results shown for DR($1/4$) in Fig. 7.16. The ghosting level is $\mu_{\ell} = 2$ for this object. In the elementary shift-and-add image the central peak has a wide base. This is to be expected since the component g, when mistaken for the reference, is shifted to the origin. The ghost to the left of the the central peak is nearly of the same amplitude as the image of g. In the image $\tilde{f}_{cor}^{(8)}$ the ghost to the left of the central peak is reduced and the image of g is enhanced. The structure of the object is shown more clearly in $\tilde{f}_{cor}^{(8)}$ than in $\tilde{f}_e^{(1)}$.

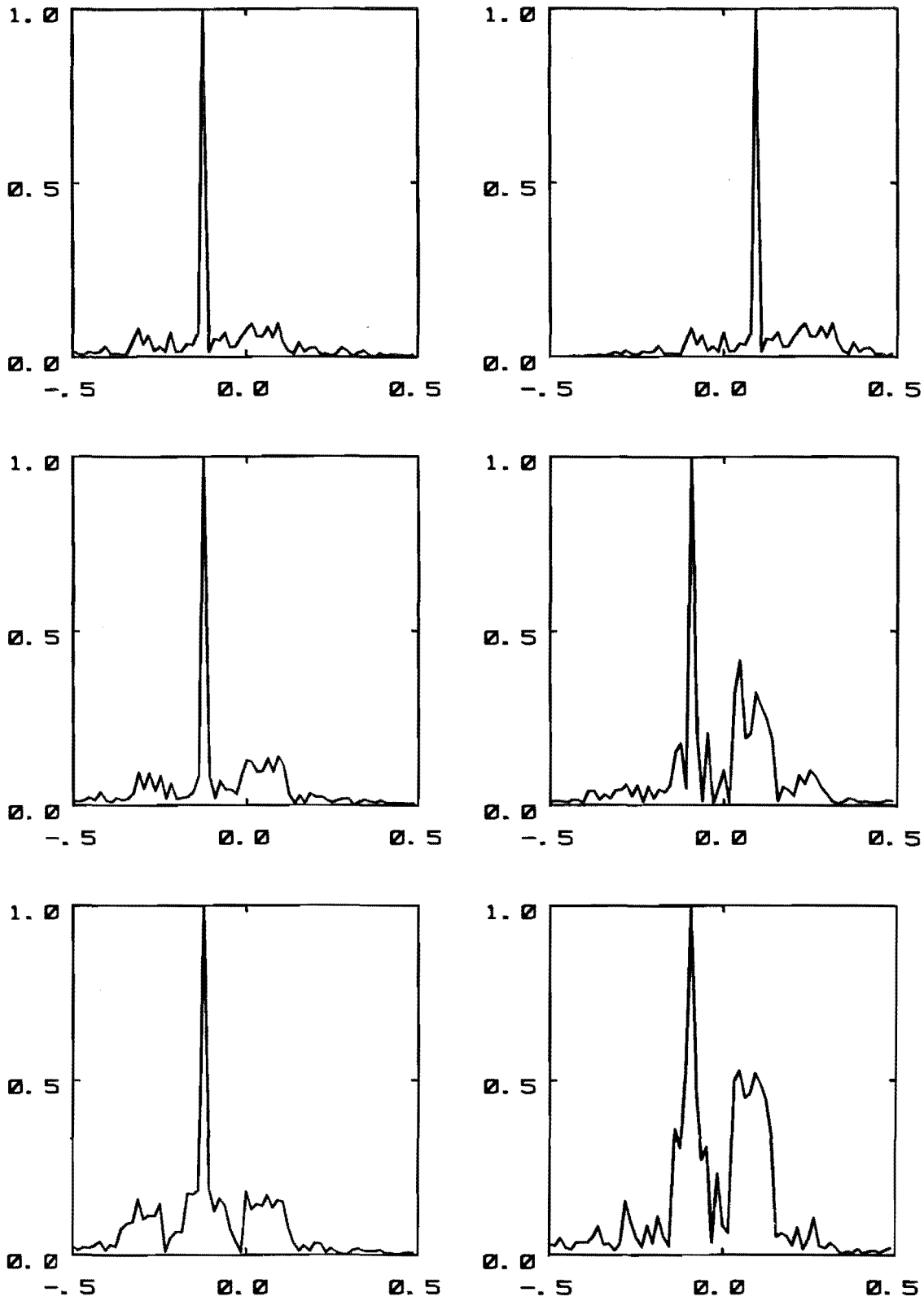


Fig. 7.16 Elementary shift-and-add and correlation shift-and-add images for the object ensemble DR(α).
Left column: $|\tilde{f}_e^{(1)}|$. Right column: $|\tilde{f}_{cor}^{(8)}|$.
Top row: DR(1/16).
Middle row: DR($\frac{1}{8}$).
Bottom row: DR($\frac{1}{4}$).

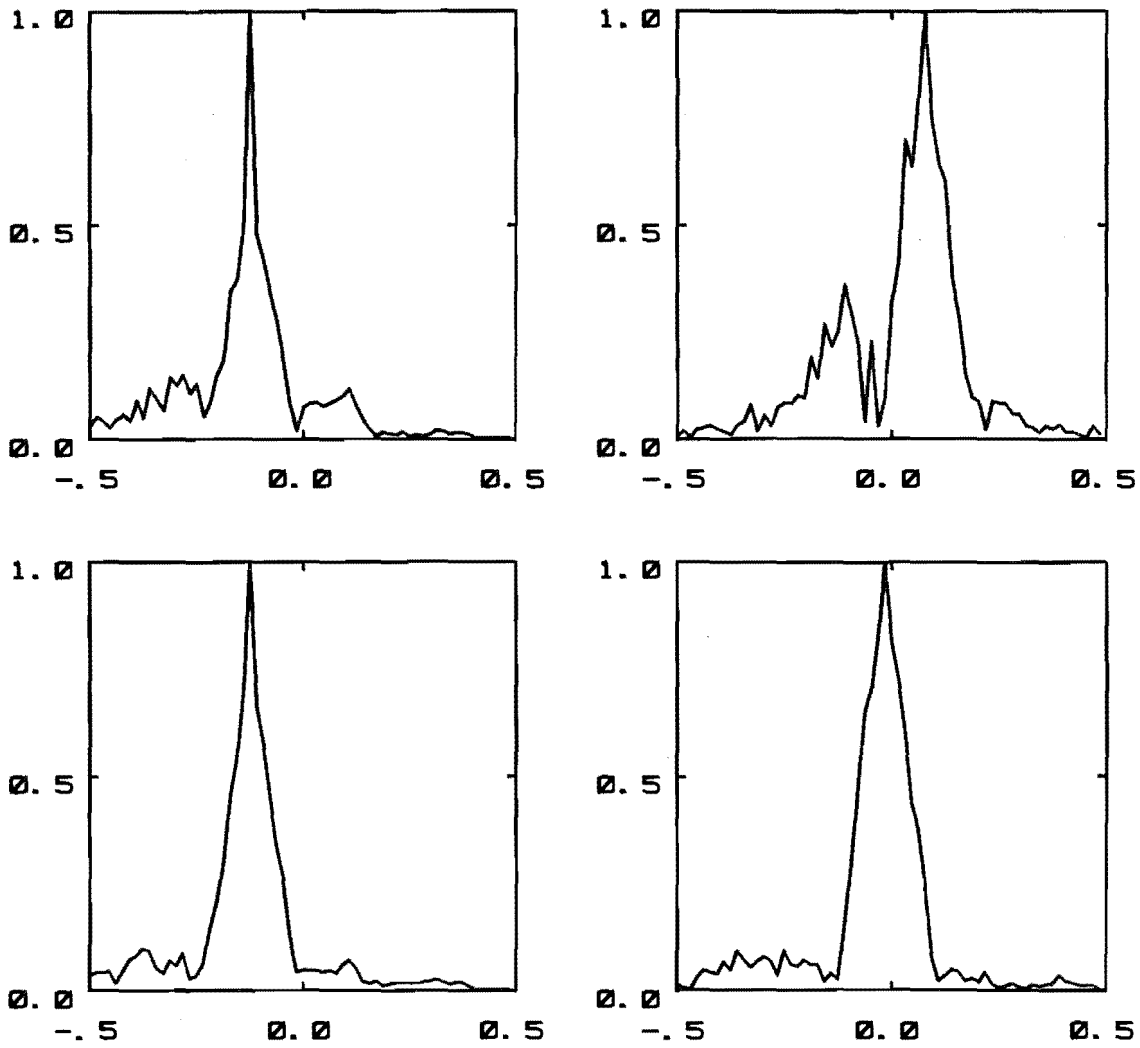


Fig. 7.16 (cont.)

Top row: DR($\frac{1}{2}$). Bottom row: DR(1).

The images of DR($\frac{1}{2}$) shown in Fig. 7.16 have quite a different appearance to that of DR($\frac{1}{4}$). The component g is now so bright ($\mu_\ell = 4$) that pixels belonging to it are chosen more often than the reference. Under these circumstances, it is not appropriate to call either of the two components on each side of the central peak in the $\tilde{f}_e^{(1)}$ image a ghost. In the $\tilde{f}_{cor}^{(8)}$ image, however, there are only two significant components. The leftmost corresponds to a heavily blurred image of the reference.

Consider the ideal effective object (¶7.2) of DR(α), which is given by the autocorrelation of (7.54), i.e.

$$f * f = \Delta^2\delta + \tfrac{1}{8}\alpha^2\, S_{1/8}(\text{tri}) + \Delta\, T_{3\hat{\mathbf{x}}/16}\left(S_{1/8}(\text{rect})\right) + \Delta\, T_{-3\hat{\mathbf{x}}/16}\left(S_{1/8}(\text{rect})\right) . \qquad (7.73)$$

For DR($\frac{1}{4}$), the ideal effective object is dominated by the triangle function in the second term on the right hand side of (7.73). A triangle is not a suitable reference for shift-and-add because it does not contain a bright isolated point. This is the reason why the image of the reference r in $\tilde{f}^{(8)}_{cor}$ is heavily blurred. The $\tilde{f}^{(8)}_{cor}$ image is superior to the $\tilde{f}^{(1)}_{e}$ image in two respects. Firstly, it indicates that the object contains two components rather than three. Secondly, the width of the stronger component (the rectangle) is reproduced more accurately. In fact, the half-amplitude width of the central peak in $\tilde{f}^{(1)}_{e}$ is 0.03 while that of the component corresponding to g in $\tilde{f}^{(8)}_{cor}$ is 0.11. The latter agrees quite well with the true width of g which is 0.125. The base of the central peak in $\tilde{f}^{(1)}_{e}$ is considerably wider than 0.03, but nevertheless the shape of the component is reproduced more faithfully in $\tilde{f}^{(8)}_{cor}$ than in $\tilde{f}^{(1)}_{e}$.

The result for DR(1) is similar to that for DR($\frac{1}{2}$). The two components are more obvious in $\tilde{f}^{(8)}_{cor}$ than in $\tilde{f}^{(1)}_{e}$. The reference component is very faint and is spread out over a region corresponding to the width of the triangle function in the ideal effective object (7.73), as is to be expected. The width of g is given more accurately by $\tilde{f}^{(8)}_{cor}$ than by $\tilde{f}^{(1)}_{e}$.

7.6.6 Three-Point Objects

Correlation shift-and-add images and elementary shift-and-add images for D13(α,β,γ), and DDD(α,β,γ) are shown in Fig. 7.17. In general, all of the images significant ghosting. Usually, at least two of the ghosts (the outer ones) can be eliminated by the extent constraint. Hence, the image $\hat{f}^{(i)}$ is usually more faithful than the corresponding $\tilde{f}^{(i)}$ image.

Consider the ideal effective object f * f for an object f consisting of N δ functions located at $\boldsymbol{\alpha}_i$,

$$f = \sum_i T_{\boldsymbol{\alpha}_i}(\delta) , \quad \boldsymbol{\alpha}_i \neq 0 \qquad (7.74)$$

and

$$f * f = \sum_i \sum_j T_{\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j}(\delta) = N\delta + \sum_{i,i\neq j} \sum_j T_{\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j}(\delta) . \qquad (7.75)$$

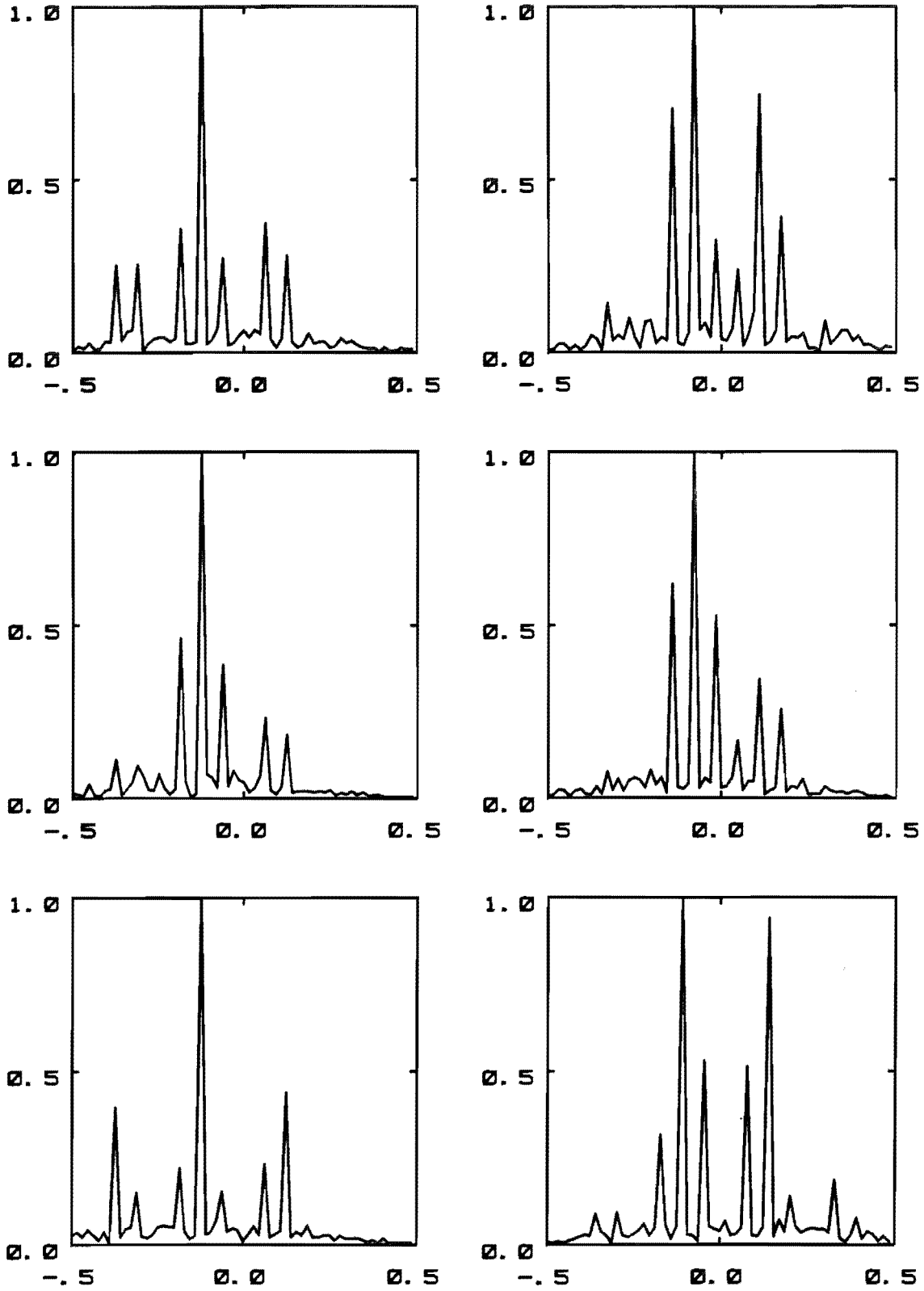


Fig. 7.17 Elementary shift-and-add and correlation shift-and-add images for three-point objects.
Left column: $|\tilde{f}_e^{(1)}|$. Right column: $|\tilde{f}_{cor}^{(8)}|$.
Top row: D13(1,1,1).
Middle row: D13(1,1,½).
Bottom row: D13(1,½,1).

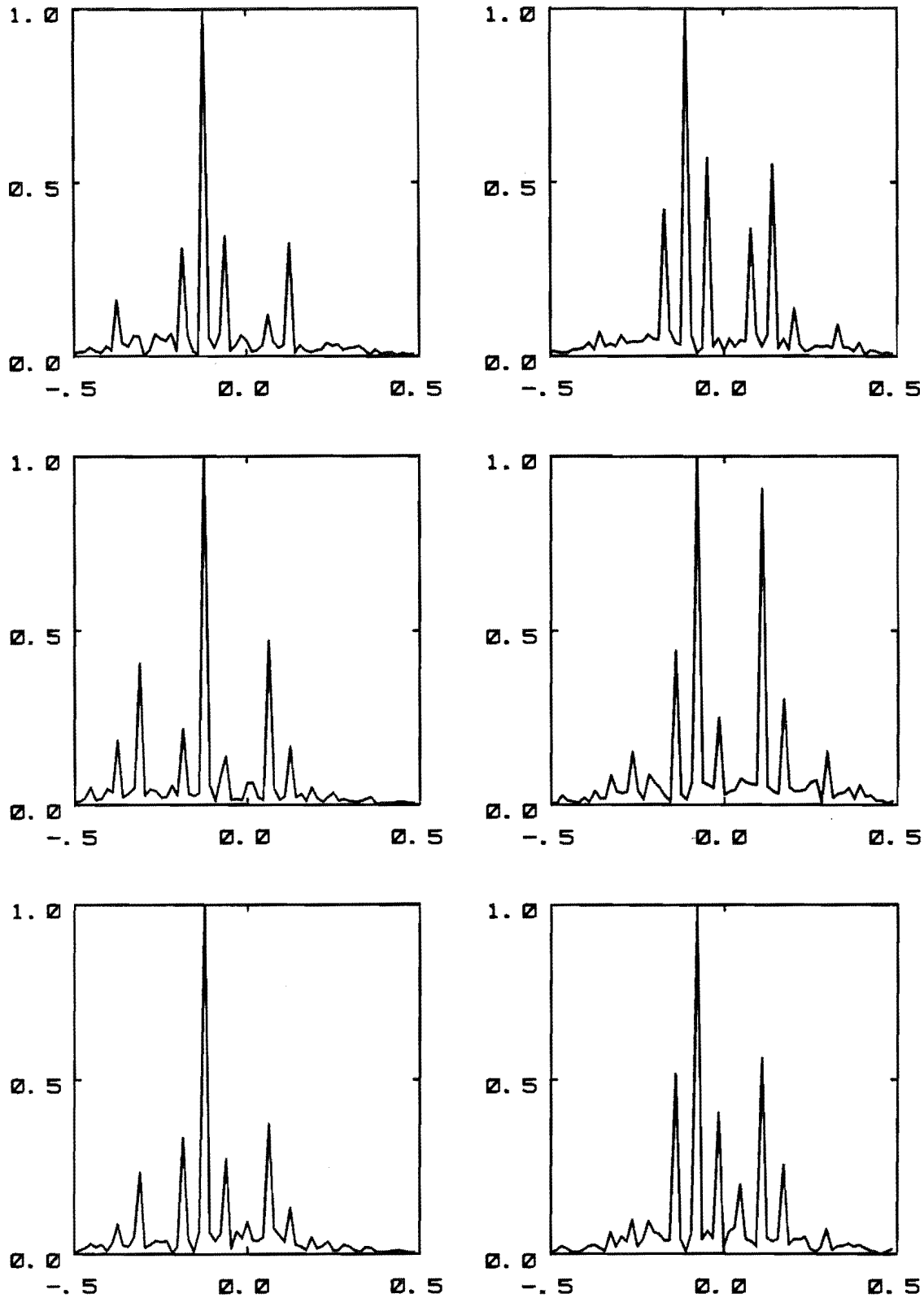


Fig. 7.17 (cont.)

Top row: D13(1, ½, ½).

Middle row: D13(½, 1, 1).

Bottom row: D13(½, 1, ½).

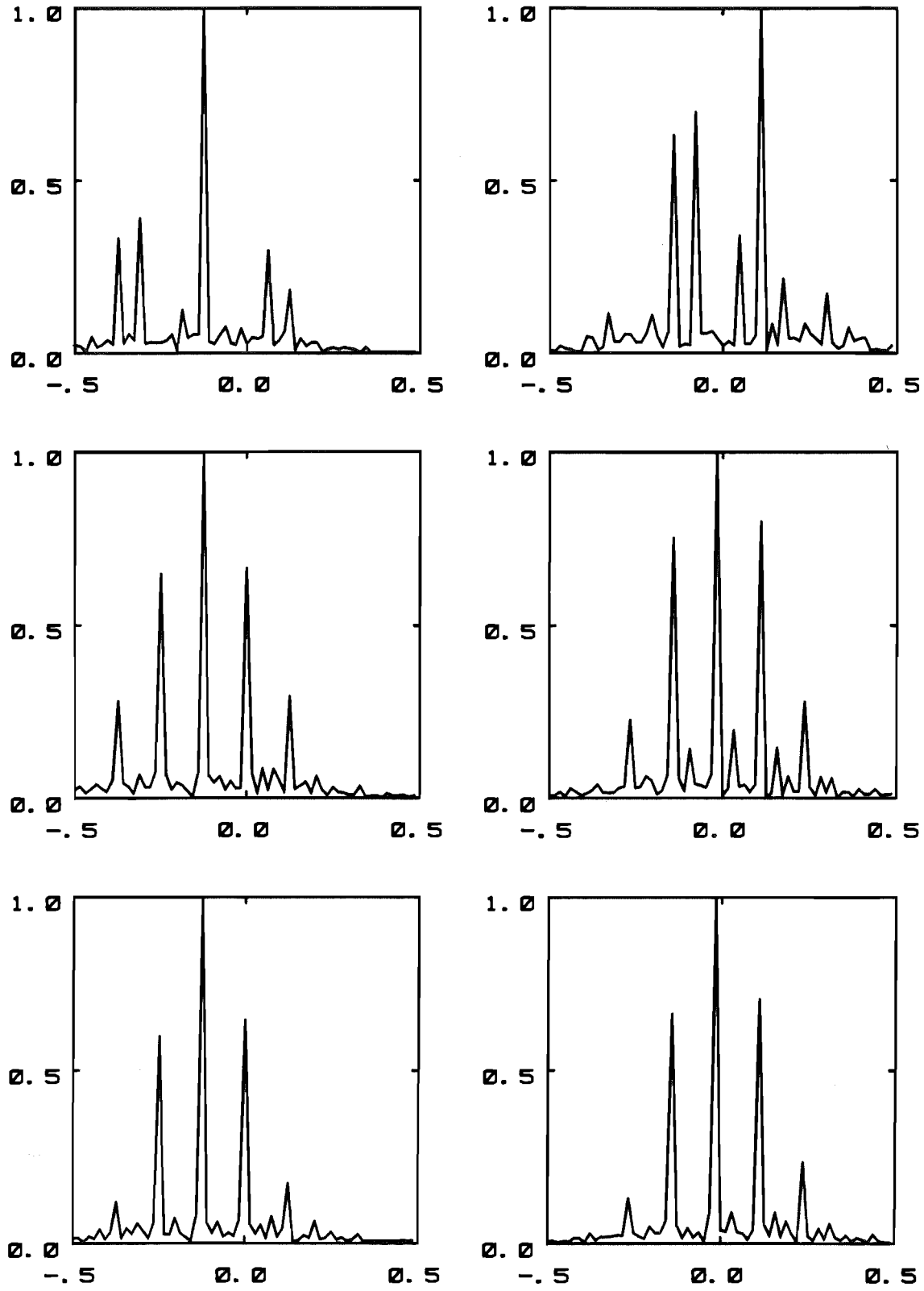


Fig. 7.17 (cont.)

Top row: D13($\frac{1}{2}$,$\frac{1}{2}$,1).

Middle row: DDD(1,1,1).

Bottom row: DDD(1,1,$\frac{1}{2}$).

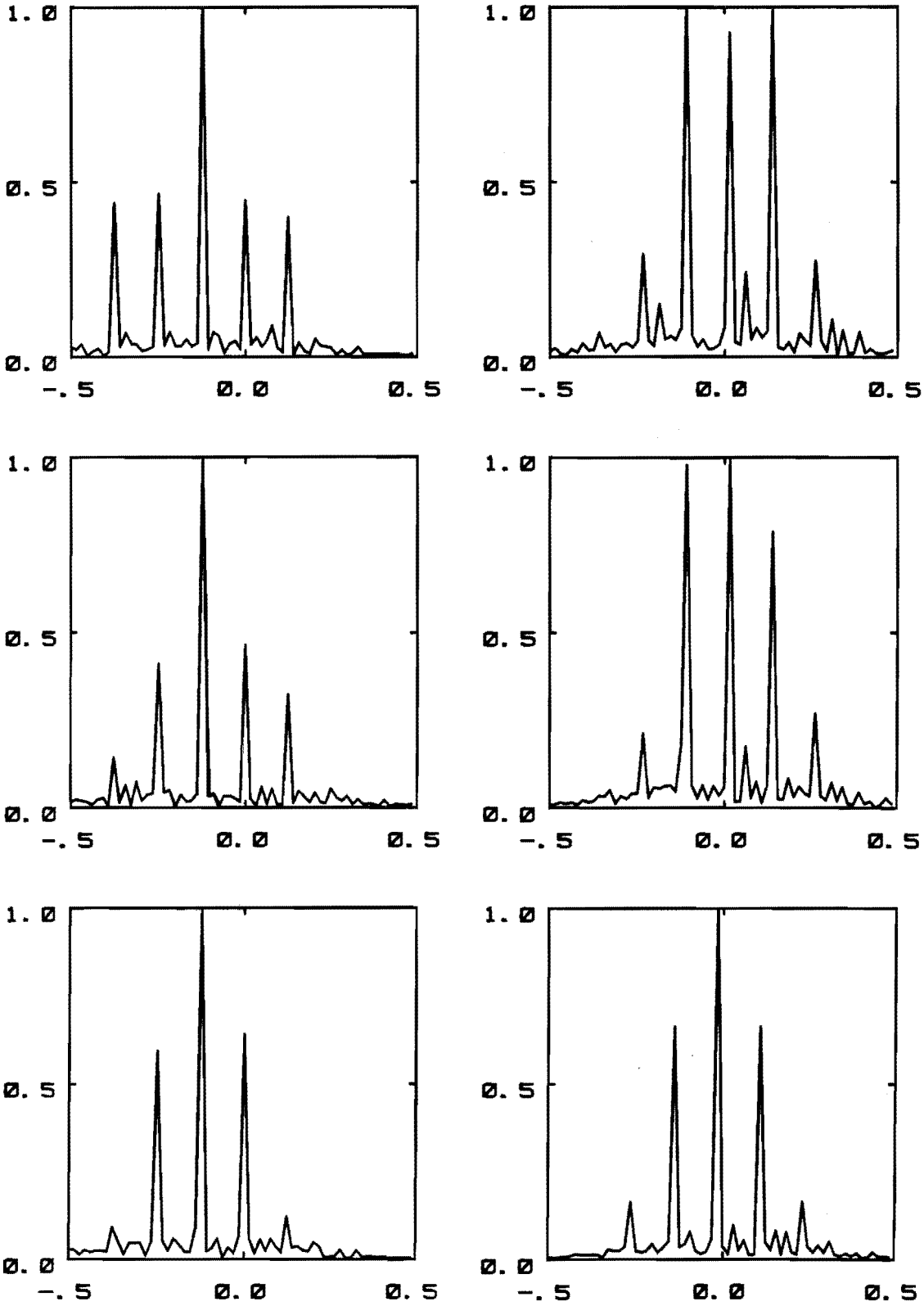


Fig. 7.17 (cont.)

Top row: DDD(1,½,1).

Middle row: DDD(1,½,½).

Bottom row: DDD(½,1,½).

The ghosting level for the ideal effective object (7.75) is obtained from (7.69) by assuming that the δ function at the origin is the reference. One has

$$\int_{T_{\mathbf{r}}} |r| \, d\mathbf{r} = N \tag{7.76}$$

and

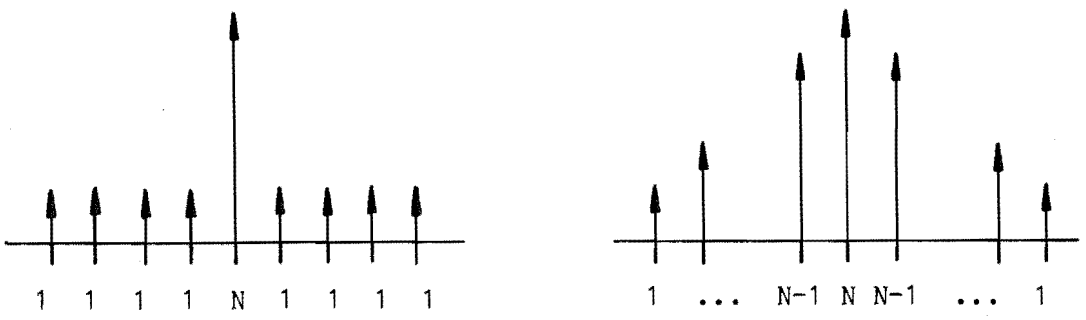
$$\int_{T_{\mathbf{r}}} \sum_{i, i \neq j} \sum_{j} T_{\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j}(\delta) \, d\mathbf{r} = N(N-1) \, . \tag{7.77}$$

Hence

$$\mu_{\ell} = N - 1 \, . \tag{7.78}$$

Note that if one of the δ functions of the object is assumed to be a reference, the ghosting level calculated from (7.69) is also N - 1. The model, upon which (7.69) is predicated, therefore predicts that the correlation shift-and-add image exhibits just as much ghosting as the ordinary shift-and-add image. However, all of the components of f have the same strength, whereas in f * f the component at the origin is stronger than the remainder. Thus, the conditions under which the model is derived (an isolated bright reference) are more nearly satisfied for f * f than for f. Since failure of the model corresponds to a high level of ghosting, the correlation shift-and-add images are likely to be better than those of elementary shift-and-add.

A similar argument indicates that the correlated shift-and-add image is more faithful if the locations $\boldsymbol{\alpha}_i$ are non-redundant. Consider Fig. 7.18 which depicts the autocorrelations of two linear arrays of δ functions, one redundant and one not. For the redundant array several of the terms in the sum on the right hand side of (7.75) have the same value for $\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j$. This results in two peaks of strength N - 1 in the autocorrelation. For the non-redundant array, all of the peaks in the autocorrelation, apart from the central one, have unit strength. Since the model (7.22) is more appropriate for the latter array, it can be expected that the correlation shift-and-add image of a non-redundant array is more faithful than that of a redundant one. The reconstructions of $D13(\alpha,\beta,\gamma)$ are in general better than those of $DDD(\alpha,\beta,\gamma)$. This is because, as mentioned in ¶7.6.3, the component spacings in $DDD(\alpha,\beta,\gamma)$ are redundant while those in $D13(\alpha,\beta,\gamma)$ are not. The images in Fig. 7.17 may be assessed by counting the number n_p of peaks greater than 0.2. Ideally, there should be three. Consider the object D13(1,1,1). In $\tilde{f}_e^{(1)}$ there are seven such



(a) (b)

Fig. 7.18 Autocorrelation of an array of δ functions.
(a) non-redundant spacings.
(b) redundant spacings.

peaks while in $\tilde{f}^{(8)}_{cor}$ there are six. Thus $\tilde{f}^{(8)}_{cor}$ is more faithful. Since the autocorrelation of D13(1,1,1) is more 'peaked' than that of D13(1,1,1), this is to be expected. For the object D13(½,1,½), $n_p = 5$ for both $\tilde{f}^{(1)}_e$ and $\tilde{f}^{(8)}_{cor}$. Since D13(½,1,½) already has a peak significantly brighter than the rest of the object, this is to be expected. For the object DDD(1,1,1), $n_p = 5$ for both $\tilde{f}^{(1)}_e$ and $\tilde{f}^{(8)}_{cor}$. Since the object is a redundant array, $\tilde{f}^{(8)}_{cor}$ is not expected to be superior to $\tilde{f}^{(1)}_e$.

Finally, the image $\tilde{f}^{(8)}_{cor}$ of DDD(1,1,1) contains one ghost which is within the region occupied by the object. Furthermore, the amplitude of the ghost is greater than the threshold level. The ghost has arisen from the contamination in the manner speculated upon in ¶7.6.4. To eliminate the ghost the threshold level would need to be raised. However, the threshold operation also removes low-level detail in the estimate corresponding to detail in the object. Thus, in general, the estimate always differs appreciably from the object. Consequently, the effective object (defined by (7.6)) does not converge to the autocorrelation but to an image resembling it.

7.6.7 Ultrasonic Objects

A grid of wires was constructed by winding plated copper wire of 0.18 mm diameter around the grid holder described in ¶5.2.3. Every third groove of the screw thread was used so the grid spacing was 1.2 mm. The wire passed through the object region eight times.

The tissue holder TH3 (¶5.2.3) was suspended on the grid. Scattering measurements were made with the holder filled with water and were repeated with the holder filled with ox liver. The measurements were made with the grid parallel to the x axis and also with the grid at 45° to the x axis.

The results of performing correlation shift-and-add on the narrow-band images are shown in Figs. 7.19 and 7.20. The grid is not recognisable in the narrow-band images. The distortions in the narrow-band images obtained through water are due to aberrations in the measurement aperture. The distortions in the narrow-band images obtained through liver are significantly worse than those obtained through water. This confirms that the liver is an inhomogeneous medium, as is of course to be expected. The elementary shift-and-add images show only two or three of the wires. However, in the correlation shift-and-add images six to nine wires are visible. The reduced amplitude of the wires at the edge of the grid when it is at 45° is probably due to the intensity of the incident beam being lower there. Thus, correlation shift-and-add is able to reconstruct significantly better images from experimental data than elementary shift-and-add.

The results presented here support the theoretical argument developed in ¶6.8 that shift-and-add is an approximate solution to the remote probing problem. Even when the grid is surrounded by ox liver the distortion introduced must have a significant isoplanatic component since faithful images are obtained by correlation shift-and-add.

7.7 SUMMARY

In this chapter the correlation shift-and-add algorithm is investigated. An analysis indicates that, for an object consisting of isolated components, the components are reconstructed in the correct locations but their amplitudes tend to be higher than in the object. Simulations performed for several classes of objects support these conclusions. Correlation shift-and-add images of an ultrasonic object

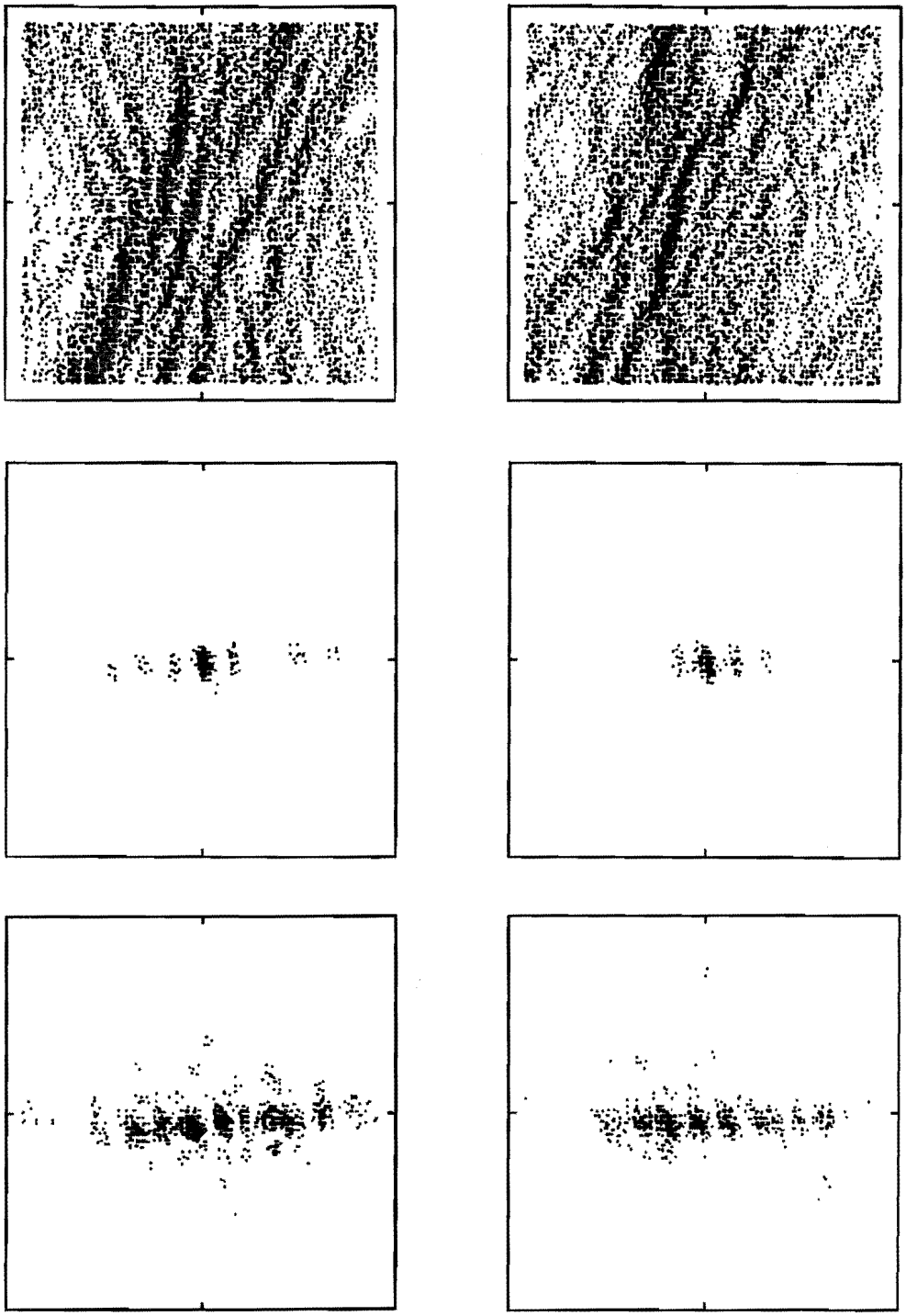


Fig. 7.19 Images of a wire grid positioned along the x axis.
Left column: TH3 filled with water.
Right column: TH3 filled with ox liver.
Top row: $|s_{75}|$. Middle row: $|\tilde{f}_e^{(1)}|$.
Bottom row: $|\tilde{f}_{cor}^{(5)}|$.
See ¶6.2 for a description of the display algorithm.
Display parameters: $d_0 = 0.1$; $d_\varepsilon = 0$; $d_1 = 1$.

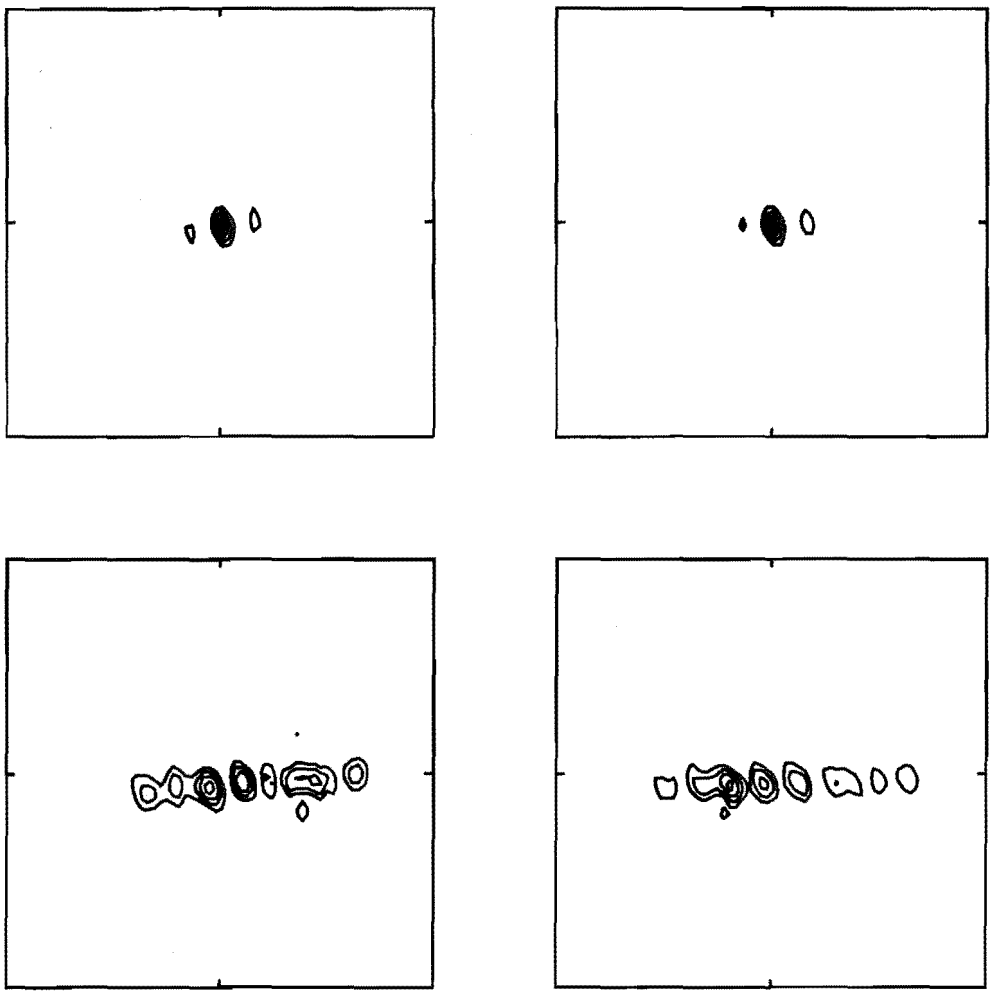


Fig. 7.19 (cont.)

Contour images. Contours at 0.3, 0.5, 0.7, 0.9.
Top row: $|\tilde{f}_e^{(1)}|$.
Bottom row: $|\tilde{f}_{cor}^{(5)}|$.

conclusions. Correlation shift-and-add images of an ultrasonic object consisting of a grid of wires are found to be superior to those generated by elementary shift-and-add.

Correlation shift-and-add is a suitable image reconstruction technique for objects dominated by one or more δ functions. In general, the images obtained are more faithful than those obtained by elementary shift-and-add. In particular, the extent of the reconstruction is closer to that of the object. Also, although ghosting artefacts are present in both correlation and elementary shift-and-add images, the ghosting artefacts are more

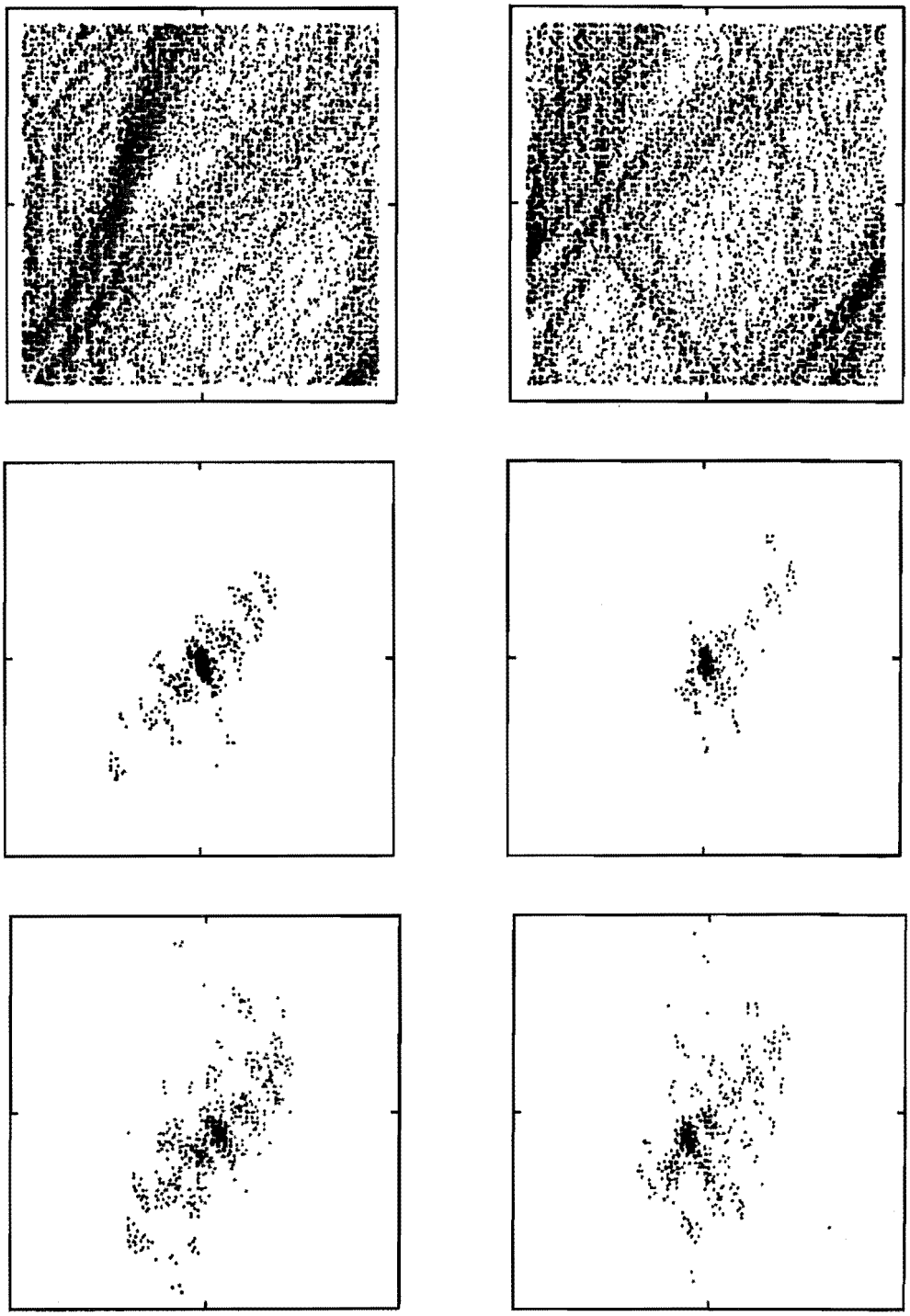


Fig. 7.20 Images of a wire grid positioned at 45° from x axis.
Left column: TH3 filled with water.
Right column: TH3 filled with ox liver.
Top row: $|s_{75}|$. Middle row: $|\tilde{f}_e^{(1)}|$.
Bottom row: $|\tilde{f}_{cor}^{(5)}|$.
Display parameters: $d_0 = 0$; $d_\varepsilon = 0$; $d_1 = 1$.

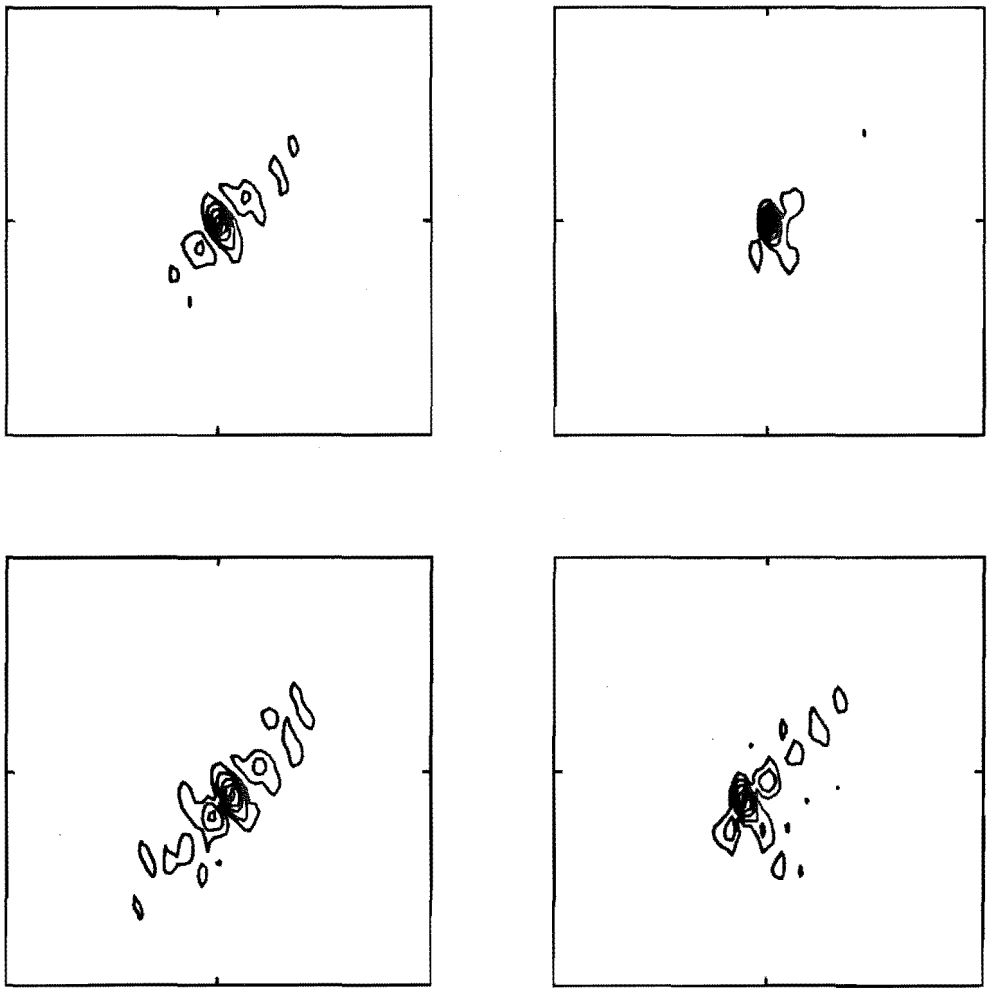


Fig. 7.20 (cont.)

Contour images.
Contours at 0.15, 0.3, 0.5, 0.7, 0.9.
Top row: $|\tilde{f}_e^{(1)}|$.
Bottom row: $|\tilde{f}_{cor}^{(5)}|$.

readily distinguished from the image of the object in the correlation shift-and-add images. In some situations, only the number of components is of interest. Since components of the object may be distinguished more readily in the image $\tilde{f}_{cor}^{(\infty)}$ than in the image $\tilde{f}_e^{(1)}$, the $\tilde{f}_{cor}^{(\infty)}$ image is more useful in this situation.

Correlation shift-and-add has two chief defects. First, for an object consisting of several components, it often does not reconstruct the amplitude of each component correctly. Second, only objects dominated by δ functions may be reconstructed. Two extensions which significantly improve upon this situation are presented in Chapter 8.

8. FURTHER EXTENSIONS TO SHIFT-AND-ADD

8.1 INTRODUCTION

The correlation technique of Chapter 7 allows objects dominated by several bright points of a similar magnitude to be reconstructed. When the amplitudes of the points differ, but are close enough to cause ghosting, the amplitudes are not well reconstructed. A deconvolution technique which attempts to overcome this deficiency and to allow extended objects to be reconstructed is evaluated in ¶8.2.

A differential approach to speckle imaging which may allow certain extended objects to be reconstructed is described and illustrated in ¶8.3.

The speckle processing discussed in ¶6.8 assumes that the Born approximation applies in the region of the object being imaged. In ¶8.4 an iterative procedure is proposed which may account for curvature of rays and variations of wave speed along them.

The results obtained in this chapter are summarised in ¶8.5.

8.2 A DECONVOLUTION TECHNIQUE

8.2.1 Method

The ideal effective object (¶7.2) for correlation shift-and-add is the autocorrelation of the object. The autocorrelations of many types of extended objects are unsuitable as shift-and-add references because they too tend to be extended. A Gaussian image provides an apposite example, since the autocorrelation of a Gaussian is another Gaussian. To see this, note that (Bracewell 1978 p.100)

$$\P F(\text{gau}) = \text{gau} \tag{8.1}$$

and hence that

$$\begin{aligned} \P F(\text{gau} * \text{gau}) &= \text{gau}^{*} . \text{gau} \\ &= S_{1/\sqrt{2}}(\text{gau}) \\ &= \P F\left[S_{\sqrt{2}}(\text{gau})/\sqrt{2}\right) , \end{aligned} \tag{8.2}$$

i.e.

$$\text{gau} * \text{gau} \qquad = S_{\sqrt{2}}(\text{gau})/\sqrt{2} \ . \tag{8.3}$$

So the autocorrelation is $\sqrt{2}$ times as wide as the object and is, in fact, a worse reference for shift-and-add than the object itself.

Since shift-and-add reconstructs a ghostless image only for an object consisting of a single δ function, the effective object should ideally be a δ function. A method called 'inverse' shift-and-add, which allows this to be achieved in principle, is now presented. Given an estimate $\hat{f}$ of the object, the speckle images s_m are transformed to $\bar{s}_m$ according to

$$\bar{s}_m = \hat{f}^{-1} \odot s_m \tag{8.4}$$

where $\hat{f}^{-1}$ is the 'convolutional inverse' of $\hat{f}$, i.e.

$$\hat{f}^{-1} \odot \hat{f} = \delta \ . \tag{8.5}$$

The existence of $\hat{f}^{-1}$ and methods for computing it are discussed in ¶8.2.3. The inverse shift-and-add image $f_{inv}[\hat{f}]$ is obtained by shifting and weighting the actual speckle images using shifts and weights determined from the corresponding transformed speckle images, i.e.

$$f_{inv} = \langle w^{(n)}\left(b(\bar{s}_m)\right)\, T_{-\mathbf{b}(\bar{s}_m)}(s_m)\rangle_m \ . \tag{8.6}$$

Substituting (6.17) into (8.4) and invoking (1.36) gives

$$\bar{s}_m = h_m \odot \bar{f} + \bar{c}_m \tag{8.7}$$

where the effective object $\bar{f}$ for inverse shift-and-add is

$$\bar{f} \ = \hat{f}^{-1} \odot f \tag{8.8}$$

and the effective contamination is

$$\bar{c}_m = \hat{f}^{-1} \odot c_m \ . \tag{8.9}$$

Thus, the ideal effective object for inverse shift-and-add which is given by (8.8) for $\hat{f} = f$ is a δ function.

Inverse shift-and-add is identical to ordinary shift-and-add when the estimate is a δ function. Consequently, an iterative scheme similar to that proposed for correlation shift-and-add in ¶7.4.2 may be constructed for inverse shift-and-add. In fact, the only equation of the scheme that needs to be modified is (7.29). The new scheme is

$$\hat{f}^{(0)} = \delta \ , \tag{8.10}$$

$$\tilde{f}^{(i)} = f_{inv}[\hat{f}^{(i-1)}] \ , \tag{8.11}$$

$$\tilde{f}_t^{(i)} = \text{thresh}(\tilde{f}^{(i)}, \varepsilon_t) \ , \tag{8.12}$$

$$c^{(i)} = e(\cdot|e_x, e_y, a_x, a_y) * |f_t^{(i)}| \ , \tag{8.13}$$

$$\hat{f}^{(i)} = e(\cdot|e_x, e_y, a_x, a_y) \ T_{-\mathbf{b}(c^{(i)})}(\tilde{f}_t^{(i)}) \ . \tag{8.14}$$

8.2.2 Implementation

Inverse shift-and-add was implemented with the aid of the cor> command module described in ¶7.3 in the following manner. Suppose the estimate for correlation shift-and-add is given by $(\hat{f}^{-1})^{\dagger}$ rather than $\hat{f}$. The transformed speckle images are given by (7.1) as

$$\bar{s}_m = (\hat{f}^{-1})^{\dagger} * s_m \ . \tag{8.15}$$

When (1.39) and (1.52) are invoked the result is identical to (8.4). Hence inverse shift-and-add may be expressed in terms of correlation shift-and-add by the formula

$$f_{inv}[\hat{f}] = f_{cor}[(\hat{f}^{-1})^{\dagger}] \ . \tag{8.16}$$

8.2.3 Deconvolution

The computation of $\hat{f}^{-1}$ is equivalent to solving the deconvolution problem (Bates and McDonnell 1985). An image $\hat{f}^{-1}$ is required which satisfies

$$\hat{f}^{-1} \odot \hat{f} = \delta \ . \tag{8.17}$$

The Fourier transform of (8.17) is

$$\hat{F}^{-1} \ \hat{F} = 1 \tag{8.18}$$

where $\hat{F}^{-1}$ and $\hat{F}$ denote the Fourier transforms of $\hat{f}^{-1}$ and $\hat{f}$ respectively. Consider (8.7). If replicas of the effective object are to be identified in the transformed speckle images, the effective contamination must be small. This places an additional constraint upon $\hat{f}^{-1}$ as is now shown.

A naive solution of (8.17) is given by

$$\hat{F}^{-1} = \frac{1}{\hat{F}} . \tag{8.19}$$

However, when $\hat{F}$ is small, $1/\hat{F}$ is large and so therefore is the Fourier transform of the effective contamination, which is given by

$$\mathcal{F}(\bar{c}_m) = C_m \hat{F}^{-1} . \tag{8.20}$$

Thus, the effective contamination is large and may swamp the isoplanatic component $h_m \odot \bar{f}$ in (8.7). A better solution is given by

$$\hat{F}^{-1} = \frac{1}{\hat{F}} \frac{\text{thresh}(\hat{F},\varepsilon_{inv})}{\hat{F}} , \tag{8.21}$$

where 'thresh' denotes one of the threshold operators defined in ¶6.5. For a discontinuous threshold (6.39), $\hat{F}^{-1}$ as given by (8.21) is zero if $|\hat{F}| < \varepsilon_{inv}$. Hence, the effective contamination does not become large. The situation for the continuous threshold (6.40) is similar. For the continuous threshold $\hat{F}^{-1}$ is given by

$$\hat{F}^{-1} = \frac{1}{\hat{F}} \frac{|\hat{F}|^{2n}}{|\hat{F}|^{2n} + \varepsilon^{2n}} . \tag{8.22}$$

When n = 1, (8.22) determines an $\hat{f}^{-1}$ called the 'Wiener filter' (Bates and McDonnell 1985 ¶16). When $n \to \infty$ the filter determined by (8.22) becomes identical to that derived from (8.21) for a discontinuous threshold. The filter derived from the discontinuous threshold is used for the simulations presented in this chapter because it is simpler to implement and its effect is more readily appreciated.

8.2.4 Analysis for Two-Point Objects

In this section the iterative scheme based upon inverse shift-and-add is shown to converge to the actual object for an arbitrary two-point object. The object has the same form as the one invoked for the analysis presented in ¶7.5, i.e.

$$f = \delta + \alpha T_{\mathbf{a}}(\delta) , \quad |\alpha| < 1 . \tag{8.23}$$

As in ¶7.5, the estimate $\hat{f}^{(i)}$ is assumed to be

$$\hat{f}^{(i)} = \delta + \beta^{(i)} T_{\alpha}(\delta) \ . \tag{8.24}$$

This formula can be justified by induction, as is now shown. The first step is to calculate the effective object $\bar{f}^{(i+1)}$. The Fourier transform of (8.24) is

$$\hat{F}^{(i)}(\mathbf{u}) = 1 + \beta \exp(-j2\pi\boldsymbol{\alpha}\cdot\mathbf{u}) \tag{8.25}$$

where $\beta^{(i)}$ is denoted by β for brevity. The magnitude of β is assumed to be less than unity so that $\hat{F}^{(i)}(\mathbf{u})$ has no zeroes. Also, the threshold level is assumed to be sufficiently small that (8.19) applies, i.e.

$$\hat{F}^{(i)-1}(\mathbf{u}) = \frac{1}{1 + \beta \exp(-j2\pi\boldsymbol{\alpha}\cdot\mathbf{u})} \tag{8.26}$$

$$= \sum_{k=0}^{\infty} (-\beta)^{k} \exp(-j2\pi k\boldsymbol{\alpha}\cdot\mathbf{u}) \tag{8.27}$$

where the binomial theorem has been invoked. Taking the inverse Fourier transform of (8.27), one obtains

$$\hat{f}^{(i)-1} = \sum_{k=0}^{\infty} (-\beta)^{k} T_{k\boldsymbol{\alpha}}(\delta) \ . \tag{8.28}$$

Note that the object may be expressed as

$$f = \hat{f}^{(i)} + (\alpha - \beta)\, T_{\boldsymbol{\alpha}}(\delta) \ . \tag{8.29}$$

This 'trick' allows the effective object, given by (8.8), to be written as

$$\bar{f}^{(i+1)} = \delta + (\alpha - \beta)\, T_{\boldsymbol{\alpha}}\left[\hat{f}^{(i)-1}\right] \ . \tag{8.30}$$

The effective object has the structure of (7.21) where

$$\bar{g} = (\alpha - \beta)\, T_{\boldsymbol{\alpha}}\left[\hat{f}^{(i)-1}\right] \ . \tag{8.31}$$

Note that by inspection of (8.31) and (8.28), the amplitude of the brightest point of $\bar{g}$ is given by

$$b(\bar{g}) = \alpha - \beta \ . \tag{8.32}$$

Hence, if the estimate is a reasonable approximation to the object then $\alpha \approx \beta$ and $\bar{g}$ is weak compared to the reference. The model (7.22) is expected to be valid under these circumstances. Before it can be applied, the integral of $|\bar{g}|$ over $T_{\mathbf{r}}$ must be evaluated, i.e.

$$\int_{T_{\mathbf{r}}} |\bar{g}| = |\alpha - \beta| \; T_{-\alpha}\Big(\int_{T_{\mathbf{r}}} |\hat{f}^{(i)-1}|\Big)$$

$$= |\alpha - \beta| \sum_{k=0}^{\infty} |\beta|^{k}$$

$$= \frac{|\alpha - \beta|}{1 - |\beta|} \tag{8.33}$$

where (8.28) and the binomial theorem are invoked. Note that

$$|\hat{f}^{(i)-1}| * f = |\hat{f}^{(i)-1\dagger}| + \alpha T_{\boldsymbol{\alpha}}(|\hat{f}^{(i)-1\dagger}|) \tag{8.34}$$

where the expression (8.23) is employed for the object. The ghosting term in the model (7.22) is given by

$$|\bar{f}^{(i+1)}| * f = f + (\alpha - \beta)[T_{-\alpha}(|\hat{f}^{(i)-1\dagger}|) + \alpha|\hat{f}^{(i)-1\dagger}|] \tag{8.35}$$

where (8.30) and (8.34) have been invoked. When (8.33) and (8.35) are substituted into the model (7.22) an expression for the inverse shift-and-add image is found, i.e.

$$\tilde{f}^{(i+1)} = \tfrac{1}{2}\Big(2 - \frac{|\alpha - \beta|}{1 - |\beta|}\Big)f + \frac{\alpha - \beta}{2}[T_{-\alpha}(|\hat{f}^{(i)-1\dagger}|) + \alpha|\hat{f}^{(i)-1\dagger}|] \; . \tag{8.36}$$

When $\beta \approx \alpha$, the second term on the right hand side of (8.36) is small. Thus, the application of the extent constraint (¶7.4.1) can be reasonably be expected to isolate the estimate $\hat{f}^{(i+1)}$ from the components of (8.36) that coincide with the first term of the right hand side of (8.36), i.e. those at **0** and $\boldsymbol{\alpha}$. Note that

$$\hat{f}^{(i)-1\dagger} = \sum_{k=0}^{\infty} (-\beta^{*})^{k} \; T_{-k\boldsymbol{\alpha}} \; (\delta) \; . \tag{8.37}$$

When (8.37) is substituted into (8.36), and components at the same location are grouped together, one obtains

$$\tilde{f}^{(i+1)} = \frac{\alpha - \beta}{2} \sum_{k=0}^{\infty} [|\beta|^{k}(1 + \alpha|\beta|) \; T_{-(k+1)\boldsymbol{\alpha}}(\delta)] + \tfrac{1}{2}\Big(\alpha(\alpha - \beta) + 2 - \frac{|\alpha - \beta|}{1 - |\beta|}\Big)\delta$$

$$+ \frac{\alpha}{2}\Big(2 \; - \frac{|\alpha - \beta|}{1 - |\beta|}\Big) \; T_{\boldsymbol{\alpha}}(\delta) \; . \tag{8.38}$$

In (8.38) one recognises an image of the object with distorted amplitudes and an array of ghosts located at $-\boldsymbol{\alpha}$ to $-\infty$ in steps of $\boldsymbol{\alpha}$. If $\beta \to \alpha$ then the ghosts disappear and the object is reconstructed faithfully, i.e. $\tilde{f} \to f$.

For the normalised estimate obtained from (8.38), $\beta^{(i+1)}$ is given by

$$\beta^{(i+1)} = \frac{\alpha}{\dfrac{\alpha(\alpha - \beta)}{2 - \dfrac{|\alpha - \beta|}{1 - |\beta|}} + 1} . \tag{8.39}$$

To simplify the algebra α and β are taken to be real and the hypothesis

$$0 \leq \beta \leq \alpha \tag{8.40}$$

is made, so that (8.39) simplifies to

$$\beta^{(i+1)} = \frac{\alpha(2 - \alpha - \beta)}{\alpha(\alpha - \beta)(1 - \beta) + 2 - \alpha - \beta} . \tag{8.41}$$

The possible limits of the sequence $\{\beta^{(i)}\}$ are found by setting

$$\beta^{(i+1)} = \beta . \tag{8.42}$$

The resulting equation is

$$(\alpha - \beta)[\alpha(\beta^2 - \beta - 1) + 2 - \beta] = 0 . \tag{8.43}$$

As expected, $\beta = \alpha$ is a possible limit, but since the term in the square brackets in (8.43) is a quadratic in β there is the possibility of two other limiting values. The discriminant Δ_d of the quadratic is given by

$$\Delta_d = (1 - \alpha)(1 - 5\alpha) . \tag{8.44}$$

Hence real solutions for β only exist for

$$\alpha \leq 1/5 \text{ or } \alpha \geq 1 . \tag{8.45}$$

For $\alpha = 1$, and $1/5$, the quadratic has the double roots 1 and 3, respectively. The former coincides with the root $\alpha = \beta$ of (8.43) while the latter is disallowed by the hypothesis (8.40). Consider the case $\alpha << 1/5$, for which the quadratic has two roots given by

$$\beta_1 = 1/\alpha - 1 - \alpha \tag{8.46}$$

and

$$\beta_2 = 2 + \alpha \tag{8.47}$$

where terms of order α^2 and higher are neglected. Both β_1 and β_2 are greater than unity and so may be disregarded in the light of (8.40) and (8.23).

Now consider the behaviour of the error $e^{(i)}$ defined by

$$e^{(i)} = \alpha - \beta^{(i)} . \tag{8.48}$$

By manipulating (8.41), one obtains

$$e^{(i+1)} = \frac{\alpha}{1 + \dfrac{2 - 2\alpha + e^{(i)}}{\alpha(1 - \alpha + e^{(i)})e^{(i)}}} . \tag{8.49}$$

Consider the case $|e^{(i)}| << 1$, with $e^{(i+1)}$ given by

$$e^{(i+1)} \approx \frac{\alpha^2 e^{(i)}}{2}\left(1 + \frac{e^{(i)}}{2\gamma} - \frac{\alpha e^{(i)}}{2}\right) \tag{8.50}$$

when $e^{(i)}/\gamma << 1$ and by

$$e^{(i+1)} \approx \alpha^2 e^{(i)}\left(1 - \frac{\gamma}{e^{(i)}} - \alpha e^{(i)}\right) \tag{8.51}$$

when $e^{(i)}/\gamma >> 1$, where

$$\gamma = 1 - \alpha . \tag{8.52}$$

When $\alpha \neq 1$, the result (8.50) shows that $\beta^{(i)} \to \alpha$ at least for $e^{(i)} << \gamma$ and, furthermore, since $e^{(i+1)}$ has the same sign as $e^{(i)}$, the hypothesis (8.40) is justified. The convergence is that of a geometric progression. The case $\alpha = 1$ is quite different, however. One has from (8.51) that

$$e^{(i+1)} \approx e^{(i)}\left(1 - e^{(i)}\right) . \tag{8.53}$$

For $e^{(i)}$ very small, the convergence is very slow. From (8.41),

$$\beta^{(i+1)} = \frac{1}{2 - \beta^{(i)}} . \tag{8.54}$$

The solution of this recurrence relation for $\beta^{(0)} = 0$ is

$$\beta^{(i)} = \frac{i}{i + 1} . \tag{8.55}$$

Hence, $\beta^{(i)} \to 1$ as required. The case $\alpha = 1$ is interesting because as $\beta^{(i)} \to 1$, the denominator of the right hand side of (8.26) tends to zero for certain $\mathbf{u}$ and thus $\hat{F}^{(i)-1} \to \infty$. In practice, the threshold operation restrains this behaviour, with the consequence that $\beta^{(i)}$ does not converge exactly to unity.

8.2.5 Results

Two-Point Objects

Iterative inverse shift-and-add has been invoked to reconstruct the objects in the DD(α) ensemble defined by (7.53). The threshold level for

the inverse filter was zero for the simulation, the results of which are shown in Fig. 8.1. The amplitudes of the components of the reconstructed objects differ from those of the actual object by less than 0.05. The ghosting artefact is just visible amongst the contamination.

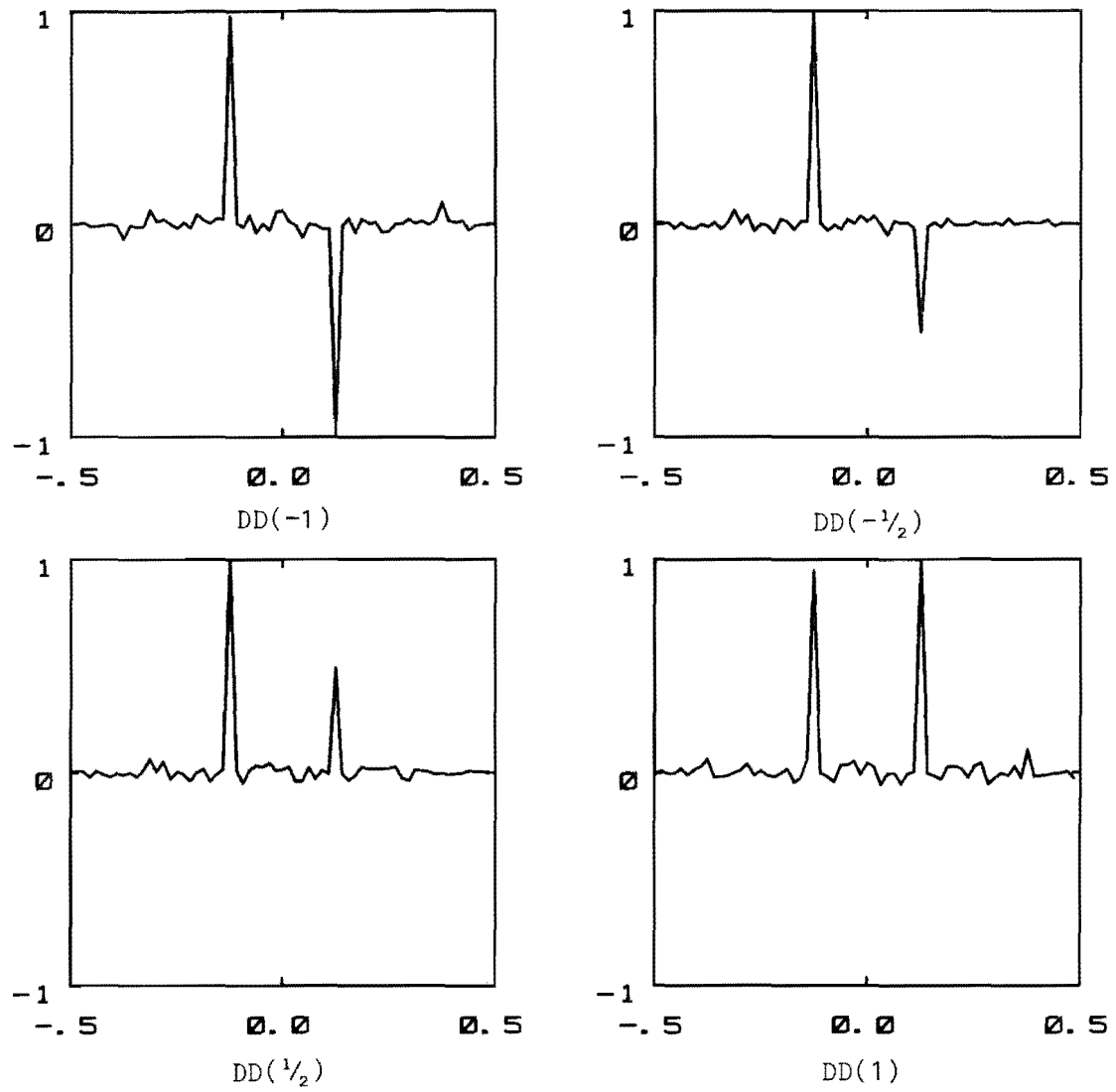


Fig. 8.1 Inverse shift-and-add images $\tilde{f}^{(8)}_{inv}$ of the object ensemble DD(α). The real parts of the images are displayed.

In general, the analysis presented in ¶8.2.4 is consistent with the results of the simulation discussed above. For instance, for the object DD(½), $e^{(1)} = 0.07$ and $e^{(2)} = 0.006$. The error is reduced by a factor of 12. The formula (8.50) predicts that it should be reduced by a factor of 8. For the object DD(1), $\beta^{(i)}$ converges to 1 in the manner indicated by (8.55),

i.e. rapidly at first and then more slowly. In most cases, the convergence of the images to the object is slightly more rapid than that expected from the model. This disagreement probably arises because the assumptions upon which the model are based are not really satisfied for the initial iterations. For these iterations $\beta^{(i)}$ is significantly different from α and the amplitude of the background in (8.30) is therefore appreciable.

Three-Point Objects

Simulations to determine the efficacy of inverse shift-and-add for reconstructing the object ensembles DDD(α,β,γ) and D13(α,β,γ) have also been carried out. The threshold used for the inverse filter defined by (8.21) is zero for the simulations presented in this subsection. This value was chosen, rather than some other arbitrary value, to simplify the interpretation of the results. The routine which computes (8.21) substitutes a zero result if the denominator is zero. However, a very small 'noisy' denominator could lead to a large noisy value for $\hat{F}^{-1}$. The simulated results were checked for such an occurrence, and none was found.

The effect of the parameters e_x and ε_t (see (8.12) and (8.13)) controlling the application of the extent constraint was determined. Three sets of simulations were carried out. For the first, $e_x = 0.4$ and $\varepsilon_t = 0.1$. For the second, $e_x = 0.4$ and $\varepsilon_t = 0.3$. For the third, $e_x = 0.7$ and $\varepsilon_t = 0.3$. The value $e_x = 0.4$ corresponds to a box just enclosing the object. The value $e_x = 0.7$ corresponds to one more than twice the linear extent of the object, i.e. effectively no extent constraint. The case of a box smaller then the object was not investigated for inverse shift-and-add, but results obtained for correlation shift-and-add suggest that the truncation caused by application of the extent constraint would prevent useful reconstructions being obtained. Certainly, an 'exact' reconstruction could not be obtained.

Since there are twelve three-point objects in the ensemble considered, and eight iterations were computed for each object, some 288 shift-and-add images $\tilde{f}^{(i)}_{inv}$ were generated. To present the results an image 'classification' scheme has therefore been devised. The results are shown in Table 8.1. The classification employed is indicated in the key for Table 8.1 and is further explained below.

Unlike correlation shift-and-add, the behaviour of the sequence of inverse shift-and-add images $\tilde{f}^{(i)}_{inv}$ often shows quite complicated features. In some cases the sequence simply does not converge and the images

Table 8.1 Inverse shift-and-add for three-point objects. The image $\tilde{f}^{(8)}_{inv}$ is classified by the error index e_3, defined in the text according to the scheme given in the key. An asterisk indicates that the result is displayed in Fig. 8.2.

Key	
Code	Interpretation
E(xact)	$0 \leqq e_3 < 0.02$
F(aithful)	$0.02 \leqq e_3 < 0.1$
≈	$0.1 \leqq e_3 < 0.2$
~	$0.2 \leqq e_3 < 0.3$
-	$0.3 \leqq e_3$
N	Not converged
O	Oscillatory
×	Divergent

Object	$e_x = 0.4$ $\varepsilon_t = 0.1$	$e_x = 0.4$ $\varepsilon_t = 0.3$	$e_x = 0.7$ $\varepsilon_t = 0.3$
DDD(1,1,1)	×	≈	E*
DDD(1,1,½)	~	F*	-
DDD(1,½,1)	F	F*	O
DDD(1,½,½)	E*	O	O
DDD(½,1,½)	~	~*	-
D13(1,1,1)	-	F*	F
D13(1,1,½)	-	F*	F
D13(1,½,1)	-	-*	N
D13(1,½,½)	F*	O	O
D13(½,1,1)	-	≈*	O
D13(½,1,½)	≈	E*	E
D13(½,½,1)	F	F*	F

In some cases the sequence simply does not converge and the images $\tilde{f}^{(i)}_{inv}$ assume a random appearance similar to that of the speckle images. This is called 'divergent' behaviour. In other cases the sequence oscillates between two 'pseudo-limits'. This is called the 'oscillatory' case. For some of the simulations it was apparent from comparing $\tilde{f}^{(7)}_{inv}$ and $\tilde{f}^{(8)}_{inv}$ that the sequence 'had not converged'. More precisely, the relative amplitudes of some of the peaks differed by more than 2% in the two images. When the sequence converges it is appropriate to calculate a figure-of-merit e_3 which is a measure of the faithfulness of the reconstruction. The figure-of-merit is defined by

$$e_3 = \frac{\max\left(\alpha - \alpha^{(8)}, \beta - \beta^{(8)}, \gamma - \gamma^{(8)}\right)}{b\left(\tilde{f}^{(8)}_{inv}\right)} \tag{8.56}$$

where $\alpha^{(i)}$, $\beta^{(i)}$, and $\gamma^{(i)}$ are the amplitudes of the peaks in $\tilde{f}^{(8)}_{inv}$ identified as corresponding to the object peaks located at $-\hat{\mathbf{x}}/8$, $\mathbf{0}$, and $\hat{\mathbf{x}}/8$ respectively. The results for which the sequence $\{\tilde{f}^{(i)}_{inv}\}$ converges are classified into five sets (each having a particular value of e_3) listed in the key for Table 8.1.

For all of the objects except DDD($\frac{1}{2}$,1,$\frac{1}{2}$) and D13(1,$\frac{1}{2}$,1), a reconstruction with $e_3 < 0.2$ is obtained. The reconstruction of D13(1,$\frac{1}{2}$,1) has e_3 in the range 0.3 to 0.4. For DDD($\frac{1}{2}$,1,$\frac{1}{2}$), e_3 is in the range 0.2 to 0.3. In general, better reconstructions are obtained when the threshold ε_t is 0.3 rather than 0.1. This is because for the higher threshold more low amplitude spurious peaks are rejected by the threshold operation. Peaks corresponding to those in the object tend to be stronger and so are not rejected. When the extent constraint is effectively removed ($e_x = 0.7$) the reconstructions are, in general, less faithful and in four of the cases were oscillatory. Two cases were oscillatory for $e_x = 0.4$ and $\varepsilon_t = 0.3$. However, the reconstruction for DDD(1,1,1) is better for $e_x = 0.7$ than for 0.4. The estimate $\hat{f}^{(2)}$ has eight peaks in the former case and four peaks in the latter. There are more peaks for $e_x = 0.7$ because the box is wider. The effective object $\bar{f}^{(3)}$ is more like a δ function for e_x than for 0.4. Hence, the inverse shift-and-add image $\tilde{f}^{(3)}_{inv}$ obtained using $\hat{f}^{(2)}$ is expected to (and does) have considerably less ghosting for $e_x = 0.7$ than for 0.4. The best reconstruction for each object is shown in Fig. 8.2. The images selected for display in Fig. 8.2 are marked by an asterisk in Table 8.1. Where possible, as is true for three quarters of the objects, the reconstruction for $e_x = 0.4$ and $\varepsilon_t = 0.3$ is chosen.

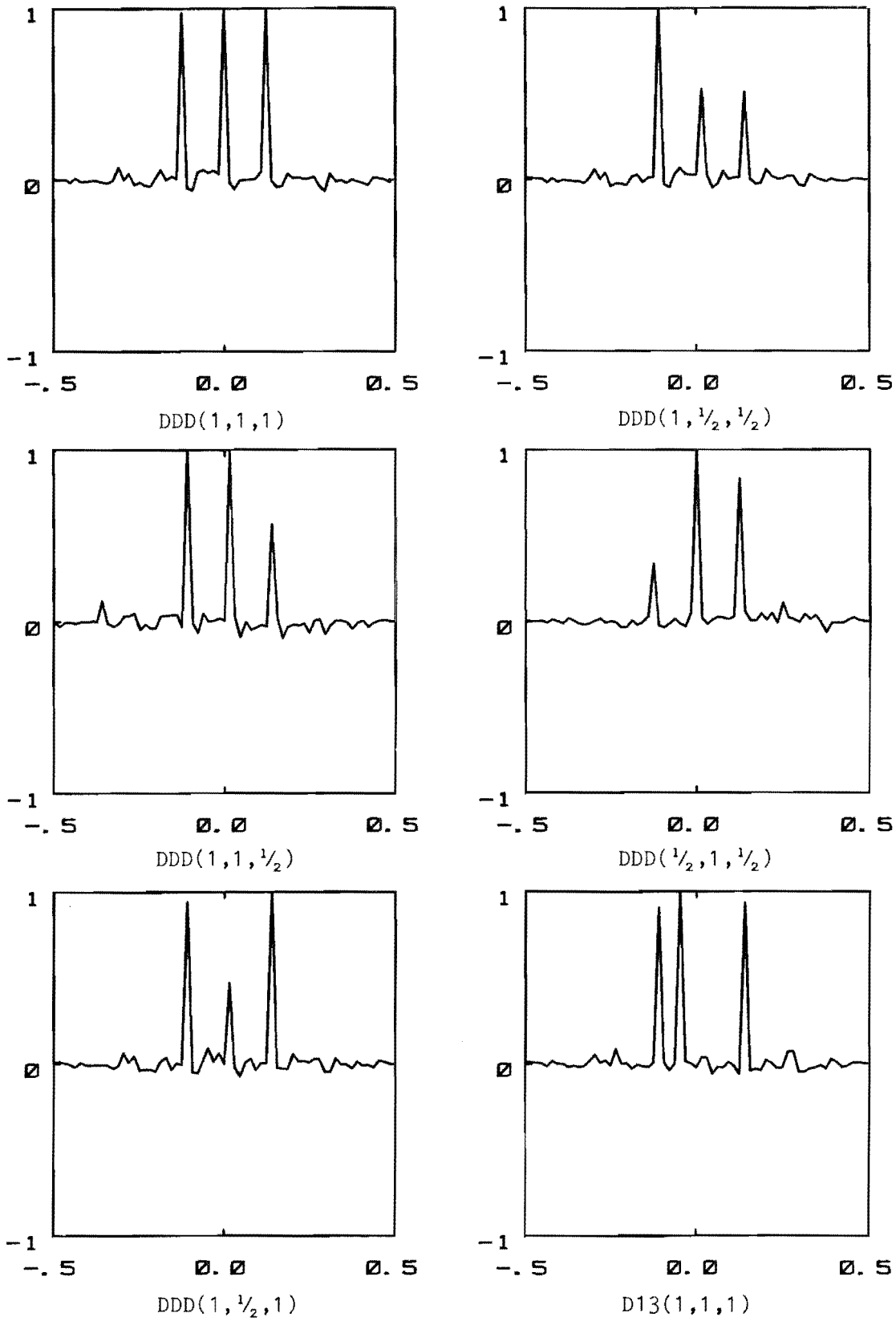


Fig. 8.2 Inverse shift-and-add images $\tilde{f}^{(8)}_{inv}$ of three-point objects. The real parts of the images are displayed.

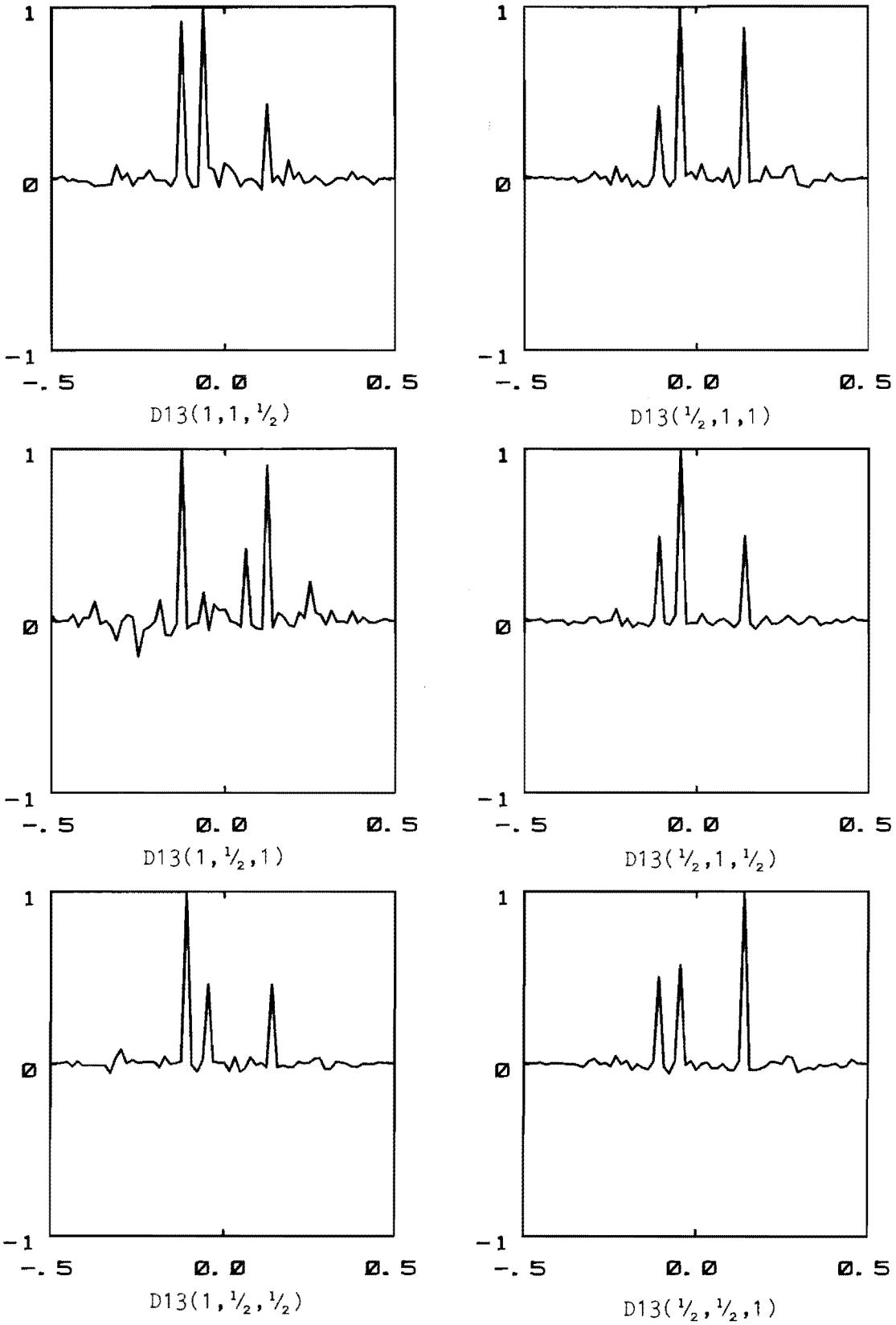


Fig. 8.2 (cont.)

The divergent behaviour for the case DDD(1,1,1), with $e_x = 0.4$ and $\varepsilon_t = 0.1$, is now discussed. The estimate $\hat{f}^{(2)}$ contains contamination from the inverse shift-and-add image $\tilde{f}^{(2)}_{inv}$. This contamination is bright enough to exceed the threshold level. The effective object $\bar{f}^{(3)}$, which is computed from the inverse of $\hat{f}^{(2)}$ and the object DDD(1,1,1), has an apparently random nature because of this. Note that the threshold level of the inverse filter (8.2) has little effect on this because it refers to a threshold operation in Fourier space rather than in image space. As a consequence of the random nature of $\bar{f}^{(3)}$, $\tilde{f}^{(3)}_{inv}$ also has a random nature and resembles a speckle image. From the third iteration onward, the shift-and-add images are speckle-like. This difficulty is overcome by raising the threshold to $\varepsilon_t = 0.3$, which removes the contamination in the estimate $\hat{f}^{(2)}$.

8.3 A DIFFERENTIAL TECHNIQUE

8.3.1 Method

In this section a method is proposed (and investigated) for reconstructing a one-dimensional piecewise constant object f from an ensemble $\{s_m\}$ of speckle images of f. Denote the derivative of f by g. The object g is an array of δ functions because f is piecewise constant. Consider the derivative s_m' of the speckle image s_m which is found from (1.47) and (6.17) to be

$$\begin{aligned} s_m' &= (h_m \odot f + c_m)' \\ &= h_m \odot f' + c_m' \\ &= h_m \odot g + c_m' \,. \end{aligned} \tag{8.57}$$

Hence, $\{s_m'\}$ is an ensemble of speckle images of g. Since g is an array of δ functions it can be approximately reconstructed by iterative correlation (or inverse) shift-and-add. Let the shift-and-add images so obtained be denoted by $\tilde{g}^{(i)}$ and the estimates by $\hat{g}^{(i)}$. From the images $\tilde{g}^{(i)}$ and $\hat{g}^{(i)}$, shift-and-add images $\tilde{f}^{(i)}$ and $\hat{f}^{(i)}$ of the object f are obtained from

$$\tilde{f}^{(i)}(x) = \int_{-\infty}^{x} \tilde{g}^{(i)}(x')\, dx' \tag{8.58}$$

and

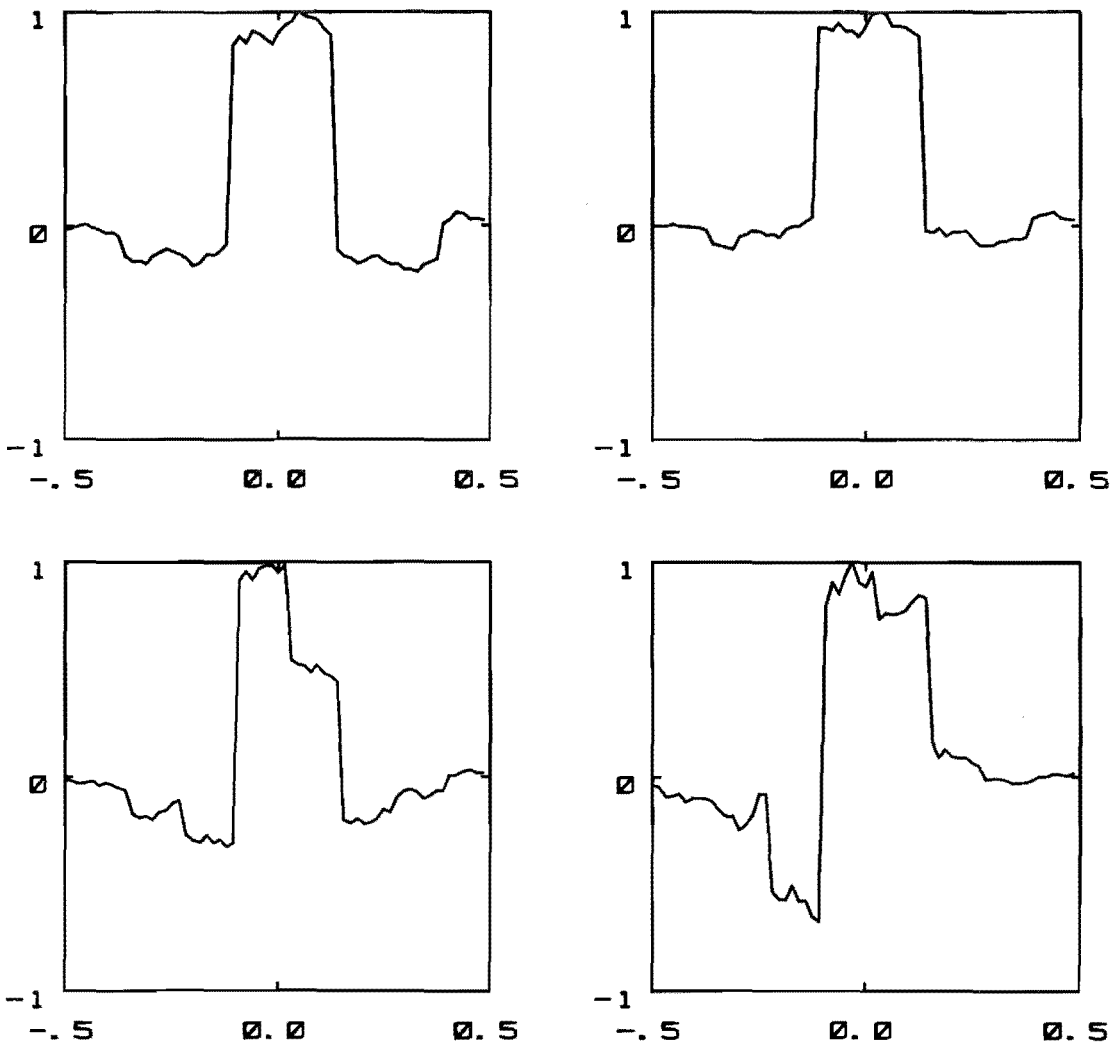


Fig. 8.3 Differential shift-and-add images of the RRE(α) object ensemble. The real parts of the images are displayed.
Left column: $\mathrm{re}\{\tilde{f}^{(8)}_{cor}\}$. Right column: $\mathrm{re}\{\tilde{f}^{(8)}_{inv}\}$.
Top row: RRE(1). Bottom row: RRE($\frac{1}{2}$).

$$\hat{f}^{(i)}(x) = \int_{-\infty}^{x} \hat{g}^{(i)}(x')\, dx' \tag{8.59}$$

where x and x' are coordinates. This method of reconstructing f is called the 'differential' technique. Essentially, it involves performing shift-and-add (or an extension of it) on the derivatives of the speckle images and then integrating the result. The image $\tilde{g}^{(i)}$ defined by (8.58) is called a 'differential' shift-and-add image.

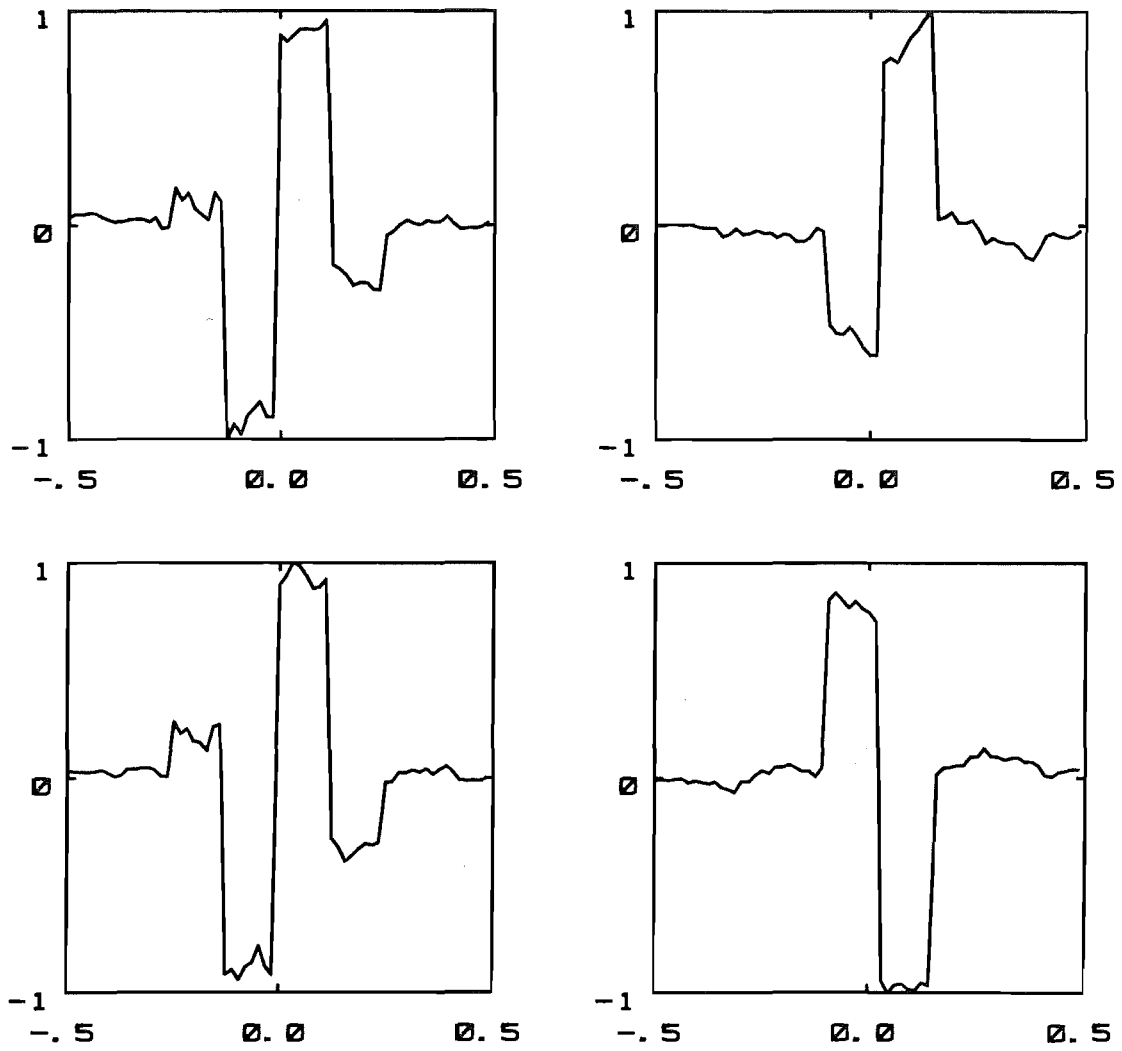


Fig. 8.3 (cont.)

Top row: RRE($-\frac{1}{2}$). Bottom row: RRE(-1).

8.3.2 Results

Simulations were carried out to test the differential technique. The object ensembles RRE(α) and RRS(α), defined by (7.55) and (7.56) respectively, were used. The ensembles of differentiated speckle images were combined by correlation and inverse shift-and-add. The results are shown in Figs. 8.3 and 8.4. The parameters (see (7.30), (7.23), and (8.21)) for the reconstructions presented in Figs. 8.3 and 8.4 are $\varepsilon_t = 0.3$; $e_x = 0.4$ and $\varepsilon_{inv} = 0$. In general, the reconstructions $\tilde{f}^{(8)}_{inv}$ obtained by inverse shift-and-add are more faithful than those obtained by correlation shift-and-add. The images of RRS(1) are an exception to this, however. The sequence $\{\tilde{f}^{(i)}_{inv}\}$ diverges and the $\tilde{f}^{(8)}_{inv}$ image shown in Fig. 8.4 has a

random structure. The reason for the divergence is similar to that discussed in ¶8.2.5. Note that some of the amplitudes of the reconstructed images are inverted with respect to those of the object. This simply reflects the fact that shift-and-add and its extensions are unable to determine the absolute amplitude of the object. For example, this applies to the $\tilde{f}^{(8)}_{inv}$ images of RRE(½) and RRS(½). Many of the images shown in Figs. 8.3 and 8.4 contain wide ghosting artefacts, such as

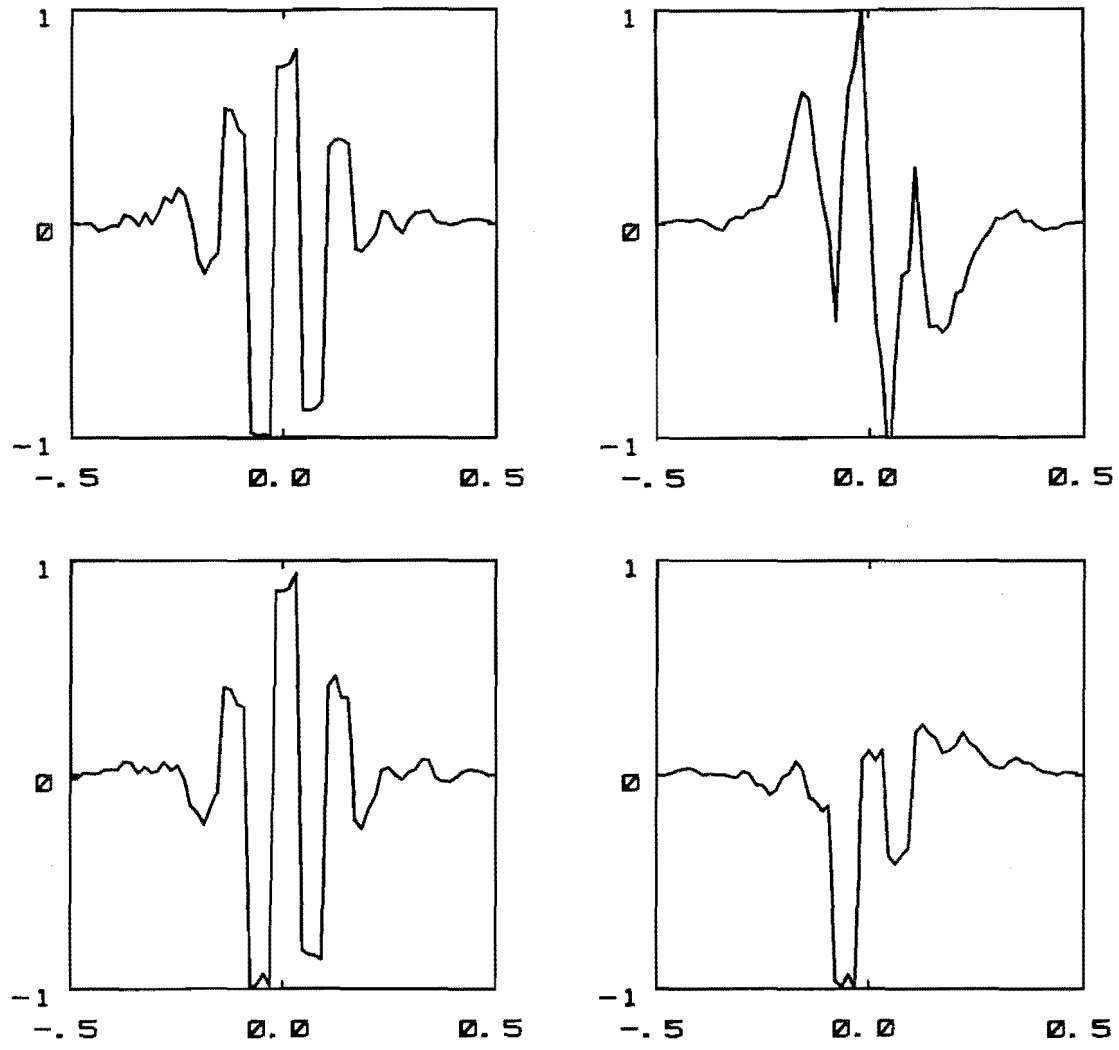


Fig. 8.4 Differential shift-and-add images of the RRS(α) object ensemble. The real parts of the images are displayed.
Left column: $re\{\tilde{f}^{(8)}_{cor}\}$. Right column: $re\{\tilde{f}^{(8)}_{inv}\}$.
Top row: RRS(1). Bottom row: RRS(½).

the $\tilde{f}^{(8)}_{inv}$ image of RRE(½) for example. This is to be expected because even a faint ghost peak in the image $\tilde{g}^{(i)}_{inv}$ (¶8.3.1) of f' is spread out

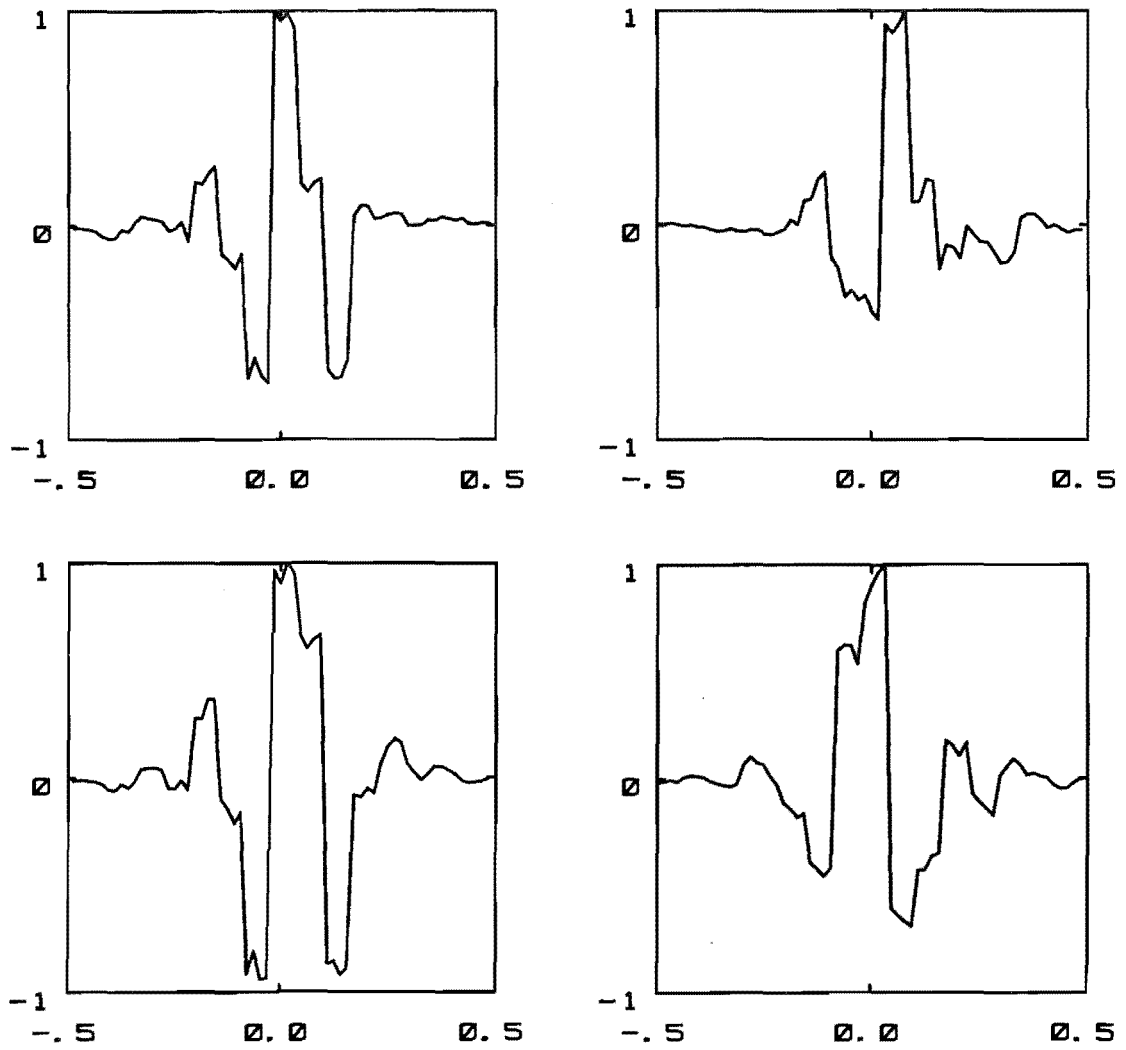


Fig. 8.4 (cont.)

Top row: RRS($-\frac{1}{2}$). Bottom row: RRS(-1).

by the integration in (8.58). The reconstructions $\tilde{f}^{(8)}_{inv}$ and $\tilde{f}^{(8)}_{cor}$ of RRE(1) are faithful. This also is to be expected since the derivative of the object is a translated version of DD(-1) and this object is well reconstructed by correlation (Fig. 7.13) and inverse (Fig. 8.1) shift-and-add.

8.4 A REFRACTION CORRECTION

When an incident wave is scattered by an object embedded in an inhomogeneous medium, there are two main refractive effects. The first is due to refraction of the incident wave as it propagates through the medium.

The wave scattered by the object also suffers refraction effects. In ¶6.8.2 it is shown that the latter effect is partially compensated for by shift-and-add. A method to correct for the former effect is proposed in this section.

In the ray coordinate formulation developed in ¶6.8.2 the incident wave ψ_{gen} is given by

$$\psi_{gen}(\bar{\mathbf{r}}) = A_{gen} \exp(-j\mathbf{k}\cdot\bar{\mathbf{r}}) \tag{8.60}$$

where $\bar{\mathbf{r}}$ is the position vector of a point in the transformed coordinate system defined by (6.113). Suppose that ψ_{gen} is propagating in the direction of the x axis in the transformed coordinate system. The effects of refraction on the phase ϕ of ψ_{gen} on the x axis in the untransformed coordinate system can be modelled by

$$\phi(x) = k_0 x + k_1 x^2 + k_2 x^3 + \ldots . \tag{8.61}$$

Suppose that a narrow-band image of the object is obtained by using the operator L defined by (6.93). The value of ϕ may be obtained wherever the magnitude of the image is significantly greater than zero. If the object consists of discrete point scatterers (as is often the case), the phase may be obtained at the location of each of the scatterers. Suppose that the phase of the image is $\theta_0, \theta_1, \ldots, \theta_N$ for the scatterers located on the x axis at points distant $x_0, x_1, \ldots, x_N$ respectively, from the origin. The measured phases θ_i are related to ϕ by

$$\theta_i = \phi(x_i) \text{ modulo } 2\pi . \tag{8.62}$$

If the scatterers are many wavelengths apart, the phase ϕ cannot be immediately determined from (8.62). Let the average of the wave number throughout the medium be denoted by $\bar{k}$. The function $(\phi(x) - \bar{k}x)$ varies considerably more slowly with x than does $\phi(x)$. Hence, it easier to obtain $\phi(x_i)$ from

$$(\theta_i - \bar{k}x_i) \text{ modulo } 2\pi = (\phi(x_i) - \bar{k}x_i) \text{ modulo } 2\pi \tag{8.63}$$

than from (8.62). From the N+1 values $\{\phi(x_i)\}$, $k_0, \ldots, k_N$ may be obtained by a polynomial fitting procedure. Now instead of assuming that the phase of the narrow-band image is $\bar{k}$, the formula (8.61) may be used. The narrow-band images are multiplied by $\exp[-j\phi(x)]$ and processed by shift-and-add. The result should be a more faithful image of the object than that obtained when the narrow-band images are multiplied by $\exp(-jkx)$.

8.5 SUMMARY

Three extensions to shift-and-add are proposed in this chapter. Each extension is intended to allow a certain class of objects to be imaged more faithfully than by ordinary shift-and-add. A refraction correction is proposed to allow objects which are embedded in a refractive medium to be better imaged. A deconvolution technique is proposed to allow arbitrary objects to be reconstructed. By simulation and analysis it is shown that the method works for two-point objects. Promising results are obtained for three-point objects by simulation. Experience with these simulations, combined with physical intuition, suggested that the method might be unsuitable for many types of extended objects. A differential technique is proposed and shown by simulation to allow a significant proportion of the class of objects which are piecewise constant to be reconstructed.

The deconvolution and differential techniques proposed in this chapter and the correlation technique investigated in Chapter 6 all employ transformed versions of the speckle images. In Chapter 9 this aspect is discussed at length and a general theory, of which the correlation, deconvolution, and differential techniques are special cases, is developed.

PART III
CONCLUSIONS

9. A GENERAL FORMULATION

9.1 INTRODUCTION

The correlation, inverse, and differential techniques described in Chapters 7 and 8 can be expressed as special cases of a more general formulation. In this formulation, shift-and-add and correlate-and-add are represented by 'averaging' operators. Such operators are discussed in ¶9.2. In the general formulation, instead of averaging the actual speckle images, an ensemble of transformed speckle images is averaged. This is called 'transform' averaging and its structure is described in ¶9.3. The special cases which correspond to the correlation, inverse, and differential techniques are identified in ¶9.4. An iterative scheme is discussed in ¶9.5. A method for combining two transformations to produce a new technique is presented in ¶9.6. A special iterative scheme is described in ¶9.7. Two classes of objects and the transform averaging techniques which allow them to be reconstructed are discussed in ¶9.8. Finally the results obtained in this thesis are discussed in ¶9.9.

9.2 DEFINITIONS

9.2.1 Averaging Operators

The 'speckle imaging' problem is to reconstruct an object f from an ensemble $\{s_m\}$ of M speckle images of the object. Each speckle image s_m is assumed to be formed according to the formula

$$s_m = f \odot h_m + c_m \ , \tag{9.1}$$

where h_m is a speckle PSF (¶6.3) and c_m is a contamination term.

Any particular algorithm for solving the speckle imaging problem can be represented by an 'averaging' operator which maps the ensemble $\{s_m\}$ to an estimate $\hat{f}$ of the object. This is written as

$$\hat{f} = A(\{s_m\}) \ . \tag{9.2}$$

It is helpful to have a notation to indicate the set of objects which an averaging operator A can reconstruct. Given an ensemble $\{s_m\}$ formed according to (9.1), the object f is reconstructed by A if

$$f = A(\{s_m\}) \ . \tag{9.3}$$

The set of objects for which (9.3) holds is denoted R(A). The equals sign in (9.3) is to be interpreted statistically, i.e. that $A(\{s_m\})$ converges to f as the number of speckle images in the ensemble increases, in the sense discussed by Middleton (1960 ¶2.1-1).

9.2.2 Shift-and-Add

Shift-and-add (Chapter 6) is one example of an averaging operator. The shift-and-add operator is defined by

$$SAA(\{s_m\})(\mathbf{r}) = \langle w(b(s_m))\ s_m(\mathbf{r} + \mathbf{b}(s_m))\rangle_m \ , \tag{9.4}$$

where $\mathbf{b}(\cdot)$ and $b(\cdot)$ are defined by (6.24) and (6.25), $\langle\cdot\rangle_m$ is defined by (6.21), and w is the weighting function given by (6.26). For the purposes of this chapter it is convenient to define two operators which are analogous to linear combinations in vector space theory (Brisley 1973 ¶8). Firstly, the 'shift-and-add combination' SAA_c (of an ensemble) is defined by

$$SAA_c(\{s_m\},\{\mathbf{r}_m\},\{w_m\})(\mathbf{r}) = \langle w_m\ s_m(\mathbf{r} + \mathbf{r}_m)\rangle_m \ , \tag{9.5}$$

where $\{\mathbf{r}_m\}$ denotes a set of M locations and $\{w_m\}$ denotes a set of M weights. The translation operator, defined by (1.31), may be used to represent the shift-and-add combination without explicitly referring to the point $\mathbf{r}$, i.e.

$$SAA_c(\{s_m\},\{\mathbf{r}_m\},\{w_m\}) = \langle w_m T_{-\mathbf{r}_m}(s_m)\rangle_m \tag{9.6}$$

$$= \langle w_m s_m \odot T_{-\mathbf{r}_m}(\delta)\rangle_m \ , \tag{9.7}$$

where (9.7) is obtained from (9.6) by invoking (1.124) and (1.45). The second operator is called the 'extended' shift-and-add combination SAA_{ec} and is defined by

$$SAA_{ec}(\{s_m\},\{t_m\}) = SAA_c(\{s_m\},\{b(t_m)\},\{w(b(t))\}) \ . \tag{9.8}$$

When (9.7) is substituted into (9.8) the formula

$$SAA_{ec}(\{s_m\},\{t_m\}) = \langle w(b(t_m))\ s_m \odot T_{-\mathbf{b}(t_m)}(\delta)\rangle \tag{9.9}$$

is obtained for SAA_{ec}. This is referred to in ¶9.2.3. The operator SAA_{ec} is similar to SAA but the locations and weights are obtained not from $\{s_m\}$ but from another ensemble $\{t_m\}$. In fact SAA is related to SAA_{ec} by

$$SAA(\{s_m\}) = SAA_{ec}(\{s_m\},\{s_m\}) \ . \tag{9.10}$$

Linearity

The shift-and-add operator is nonlinear in the sense that, in general,

$$SAA(\{t \odot s_m\}) \neq t \odot SAA(\{s_m\}) \ . \tag{9.11}$$

However, SAA_c and SAA_{ec} are linear with respect to $\{s_m\}$. This is seen by applying the identity (1.36) to (9.7), i.e.

$$\begin{aligned} SAA_c(\{t \odot s_m\},\{r_m\},\{w_m\}) &= \langle w_m(t \odot s_m) \odot T_{-r_m}(\delta)\rangle_m \\ &= t \odot \langle w_m s_m \odot T_{-r_m}(\delta)\rangle_m \\ &= t \odot SAA_c(\{s_m\},\{r_m\},\{w_m\}) \ . \end{aligned} \tag{9.12}$$

The linearity of SAA_{ec} with respect to $\{s_m\}$ follows from (9.12) and (9.8), i.e.

$$\begin{aligned} SAA_{ec}(\{t \odot s_m\},\{t_m\}) &= SAA_c(\{t \odot s_m\},\{b(t_m)\},\{w(b(t_m))\}) \\ &= t \odot SAA_c(\{s_m\},\{b(t_m)\},\{w(b(t_m))\}) \\ &= t \odot SAA_{ec}(\{s_m\},\{t_m\}) \ . \end{aligned} \tag{9.13}$$

The extended shift-and-add combination is nonlinear with respect to $\{t_m\}$.

9.2.3 Correlate-and-Add

A second example of an averaging operator is provided by the correlate-and-add algorithm defined in ¶6.5. The 'correlate-and-add' operator CAA is defined by

$$CAA(\{s_m\}) = \langle thresh(W(s_m),\varepsilon_m) * s_m\rangle_m \ , \tag{9.14}$$

where each ε_m is a threshold level and W is an image weighting operator. The threshold image operator 'thresh' and the image weighting operator are discussed in ¶6.5 and ¶6.4.4 respectively. A correlate-and-add extended combination operator CAA_{ec} can be defined in an analogous manner to SAA_{ec}, i.e.

$$CAA_{ec}(\{s_m\},\{t_m\}) = \langle thresh(W(t_m),\varepsilon_m) * s_m\rangle_m \ . \tag{9.15}$$

CAA and CAA_{ec} are related by

$$CAA(\{s_m\}) = CAA_{ec}(\{s_m\},\{s_m\}) \ . \tag{9.16}$$

In certain circumstances correlate-and-add gives the same result as shift-and-add. These circumstances are: (i) The image weighting function for correlate-and-add can be expressed in terms of the weighting function for shift-and-add by

$$W(t_m)(\mathbf{r}) = w^*(t_m(\mathbf{r})) \ ; \tag{9.17}$$

(ii) The threshold levels are chosen to be

$$\varepsilon_m = W(t_m)(\mathbf{b}(t_m)) \ ; \tag{9.18}$$

(iii) For each t_m there is only one point $\mathbf{r}$ for which $|t_m(\mathbf{r})|$ attains its maximum value, i.e. $\mathbf{b}(t_m)$ is unique. Taken together, (i), (ii) and (iii) imply that

$$thresh(W(t_m),\varepsilon_m)(\mathbf{r}) = w^*(b(t_m))\delta(\mathbf{r} - \mathbf{b}(t_m)) \tag{9.19}$$

or

$$thresh(W(t_m),\varepsilon_m) = w^*(b(t_m))T_{\mathbf{b}(t_m)}(\delta) \ . \tag{9.20}$$

When (9.20) is substituted into the definition of the extended correlate-and-add combination (9.15) and the identities (1.39) and (1.53) are invoked, an extended shift-and-add combination having the form of (9.9) is obtained,

$$\begin{aligned} CAA_{ec}(\{s_m\},\{t_m\}) &= \langle w^*(b(t_m))\ T_{\mathbf{b}(t_m)}(\delta) * s_m\rangle_m \\ &= \langle [w^*(b(t_m))\ T_{\mathbf{b}(t_m)}(\delta)]^\dagger \odot s_m\rangle_m \\ &= \langle w(b(t_m))\ T_{-\mathbf{b}(t_m)}(\delta) \odot s_m\rangle_m \\ &= SAA_{ec}(\{s_m\},\{t_m\}) \ . \end{aligned} \tag{9.21}$$

Linearity

CAA_{ec} is nonlinear with respect to $\{t_m\}$ but linear with respect to $\{s_m\}$. To see this, (1.40) is applied to (9.15) giving

$$CAA_{ec}(\{t \odot s_m\},\{t_m\}) = \langle thresh(W(t_m),\varepsilon_m) * (t \odot s_m)\rangle_m$$

$$= t \odot \langle \text{thresh}(W(t_m), \varepsilon_m) * s_m \rangle_m$$

$$= t \odot CAA_{ec}(\{s_m\}, \{t_m\}) \ . \qquad (9.22)$$

CAA is nonlinear. The significance of this is demonstrated by considering the image $\tilde{f}$ obtained from $\{s_m\}$ according to

$$\tilde{f} = t^{-1} \odot CAA(\{t \odot s_m\}) \ , \qquad (9.23)$$

where t is an arbitrary function and t^{-1} is its convolutional inverse (defined in ¶8.2.1). If CAA is linear, (9.23) simplifies to

$$\tilde{f} = CAA(\{s_m\}) \ . \qquad (9.24)$$

Since CAA is actually nonlinear, the image $\tilde{f}$ given by (9.23) differs from that given by (9.24). Hence, there is the possibility that by choosing t appropriately, $\tilde{f}$ can be made a more faithful image of f.

9.2.4 Extended Combinations

An extended combination is defined for SAA by (9.8) and for CAA by (9.15). Both extended combinations satisfy (see (9.10) and (9.16))

$$A(\{s_m\}) = A_{ec}(\{s_m\}, \{s_m\}) \ , \qquad (9.25)$$

where A is SAA or CAA. Also, both extended combinations are linear with respect to the ensemble to be averaged. This is expressed by (9.13) and (9.22). The two formulas may be combined by writing

$$A_{ec}(\{t \odot s_m\}, \{t_m\}) = t \odot A_{ec}(\{s_m\}, \{t_m\}) \ , \qquad (9.26)$$

where A_{ec} is either SAA_{ec} or CAA_{ec}.

9.3 A TRANSFORM TECHNIQUE

As shown in ¶6.4, the shift-and-add operator reconstructs ghostless (¶6.7) images for only a small class of objects. The objects which make up this class are zero everywhere except for a single peak.

A formal procedure for expanding the set of objects which can be reconstructed by an averaging operator A is shown in Fig. 9.1. The input to the procedure is an ensemble of speckle images $\{s_m\}$ of an object f. Each speckle image is formed according to (9.1). A transformed version $\bar{s}$ of each speckle image is obtained by applying an invertible operator T:

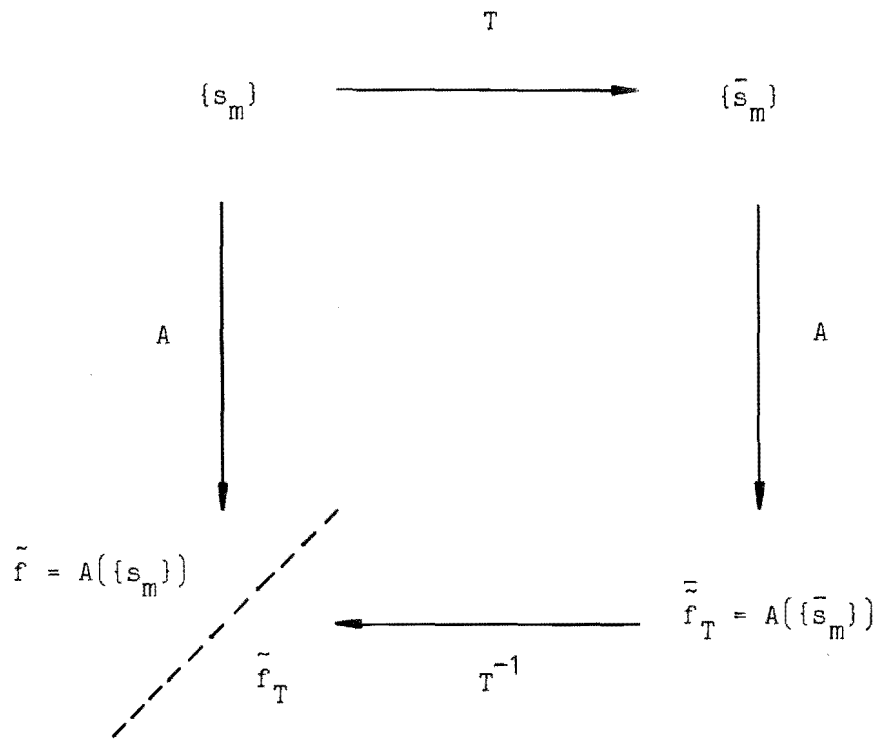


Fig. 9.1 The transformation technique. In general $\tilde{f}$ and $\tilde{f}_T$ are different.

$$\bar{s}_m = T(s_m) \ . \tag{9.27}$$

The ensemble $\{s_m\}$ is referred to as the actual ensemble in contradistinction to the transformed ensemble $\{\bar{s}_m\}$. The result of averaging the transformed ensemble using A is denoted $\bar{\tilde{f}}_T$, i.e.

$$\bar{\tilde{f}}_T = A(\{\bar{s}_m\}) \ . \tag{9.28}$$

An estimate of the object $\tilde{f}_T$ is obtained by applying the inverse of T to $\bar{\tilde{f}}_T$. This is written as

$$\tilde{f}_T = T^{-1}(\bar{\tilde{f}}_T) \ , \tag{9.29}$$

where T^{-1} is the inverse of T. The result of averaging the actual ensemble is written as

$$\tilde{f} = A(\{s_m\}) \ . \tag{9.30}$$

Steps (9.27), (9.28) and (9.29) may be condensed to

$$\tilde{f}_T = T^{-1}[A(\{T(s_m)\})] \ . \tag{9.31}$$

It is convenient to define an averaging operator A_T which maps the ensemble $\{s_m\}$ to $\tilde{f}_T$ for a given transformation T by

$$\tilde{f}_T = A_T(T,\{s_m\}) \ . \tag{9.32}$$

A_T is called the '(T-)transform' operator corresponding to A. The formula for constructing A_T from A is

$$A_T(T,\{s_m\}) = T^{-1}[A(\{T(s_m)\})] \ . \tag{9.33}$$

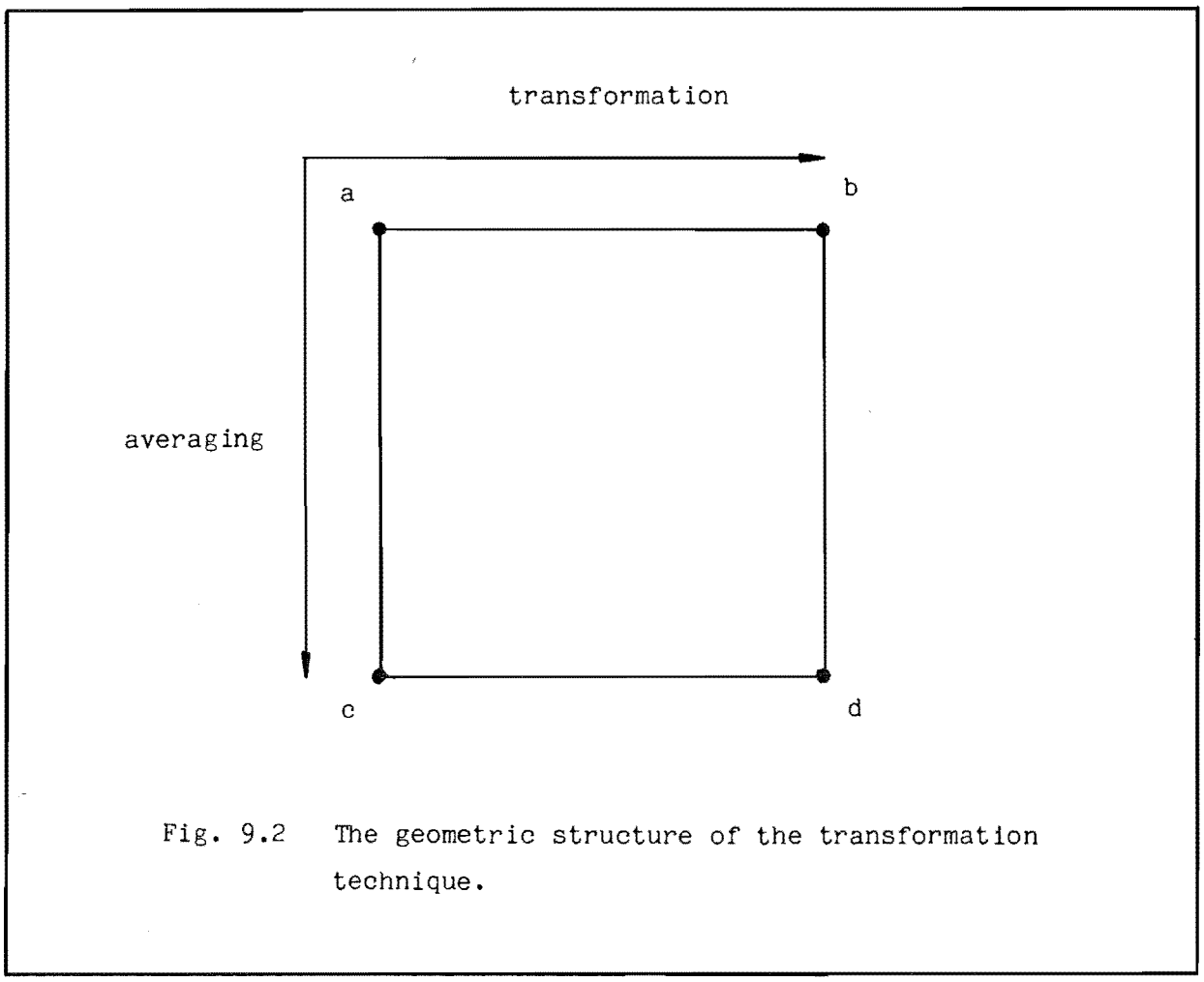


Fig. 9.2 The geometric structure of the transformation technique.

Fig. 9.1 represents the structure of the transformation procedure geometrically. To further illustrate this, consider the square depicted in Fig. 9.2. The sides and the corners of the square represent the entities used to define the transformation procedure. A corner represents either an image or a set of images. For instance. 'a' represents $\{s_m\}$, 'b' represents $\{\bar{s}_m\}$, 'c' represents $\bar{f}_T$, and 'd' represents $\tilde{f}$ and $\tilde{f}_T$. A side represents an operation. Moving to the left corresponds to the transformation T. Moving to the right corresponds to T^{-1}. Moving downwards represents the averaging process A. There are two ways to move from an upper corner to a lower corner along the sides of the square: clockwise, and anticlockwise. The two ways are equivalent if the results of carrying out the operations associated with each path are identical. If the two ways are equivalent the diagram is said to commute. For averaging operators such as SAA and CAA, the diagram does not commute because of the complicated dependence of the reconstructed image upon the object. For instance the images $\tilde{f}$ and $\tilde{f}_T$, obtained by moving from 'a' to 'c' directly, and indirectly via 'b' and 'd', are not usually the same. In Fig. 9.1 they are shown separated by a dashed line to indicate this.

The transformation approach is useful if a transformation T can be found for which $\tilde{f}_T$ is a better estimate of f than $\tilde{f}$ is. The approach is more useful still if the same transformation is suitable for a large class of objects. Ideally, one would like to find a single transformation T suitable for all objects.

9.4 CONVOLUTION AVERAGING

9.4.1 General Considerations

This section is concerned with the class of transformations for which T can be written as

$$T(g) = t \circledcirc g \ , \tag{9.34}$$

where t is a given 'filter' function. The averaging operator which maps $\{s_m\}$ to $\tilde{f}_T$ for a given filter function t is written as $A_{conv}(t,\cdot)$ and is called the 'convolution' averaging operator corresponding to A with filter function t, i.e.

$$\tilde{f}_T = A_{conv}(t,\{s_m\}) \ . \tag{9.35}$$

A formula for A_{conv} is obtained by substituting (9.35) and (9.34) into

(9.33), i.e.

$$A_{conv}(t,\{s_m\}) = t^{-1} \odot A(\{t \odot s_m\}) , \tag{9.36}$$

where t^{-1} denotes the convolutional inverse (¶8.2.1) of t.

The convolution averaging operators corresponding to SAA and CAA are written as SAA_{conv} and CAA_{conv} respectively. SAA_{conv} is called the 'convolution shift-and-add' operator and similarly CAA_{conv} is called the 'convolution correlate-and-add' operator. Let A_{conv} denote SAA_{conv} or CAA_{conv}. Expressing A in terms of A_{ec} with the aid of (9.25) and then invoking (9.26),

$$\begin{aligned} A_{conv}(t,\{s_m\}) &= t^{-1} \odot A(\{t \odot s_m\}) \\ &= t^{-1} \odot A_{ec}(\{t \odot s_m\},\{t \odot s_m\}) \\ &= (t^{-1} \odot t) \odot A_{ec}(\{s_m\},\{t \odot s_m\}) \\ &= \delta \odot A_{ec}(\{s_m\},\{t \odot s_m\}) \\ &= A_{ec}(\{s_m\},\{t \odot s_m\}) . \end{aligned} \tag{9.37}$$

So both SAA_{conv} and CAA_{conv} may be evaluated without calculating the inverse of any transformations, contrary to what one would expect from (9.36). This result shows that convolution shift-and-add, for example, may be implemented by forming a shift-and-add combination of the actual speckle images with the locations and weights derived from the brightest points of the transformed speckle images.

9.4.2 The Effective Object and the Effective Contamination

Upon substituting (9.1) into (9.27) and using (9.34), the following relationship between the object and the transformed speckle images is obtained,

$$\tilde{s}_m = t \odot (f \odot h_m + c_m) . \tag{9.38}$$

This may be rearranged, using the identities (1.36) and (1.37), to give

$$\tilde{s}_m = (t \odot f) \odot h_m + t \odot c_m . \tag{9.39}$$

Comparing (9.39) and (9.1), it can be seen that the transformed ensemble $\{\tilde{s}_m\}$ can be regarded as a set of speckle images of an object $t \odot f$ with

contaminations $t \odot c_m$. It therefore makes sense to define an effective object $\bar{f}$ by

$$\bar{f} = t \odot f \tag{9.40}$$

and effective contaminations by

$$\bar{c}_m = t \odot c_m \ , \ m = 1, \ldots, M \ . \tag{9.41}$$

These are the transformed versions of the object and the contaminations respectively. With the aid of (9.40) and (9.41), (9.39) may be written as

$$\bar{s}_m = \bar{f} \odot h_m + \bar{c}_m \ . \tag{9.42}$$

The effective object and contamination determine how successful $A_{conv}(t,\cdot)$ is at reconstructing the object. For instance, suppose that the effective object has a single bright peak and that the actual object does not. Furthermore, suppose that the effective contaminations and actual contaminations are similar. In this case, SAA_{conv} is expected to reconstruct the object more faithfully than SAA.

The assertion of the previous paragraph can be expressed as a theorem: an object f can be reconstructed by a convolution averaging operator A_{conv}, i.e.

$$f \text{ is a member of } R(A_{conv}), \tag{9.43}$$

if and only if the effective object $\bar{f}$ can be reconstructed by the operator A, i.e.

$$\bar{f} \text{ is a member of } R(A) \ . \tag{9.44}$$

To prove that (9.44) implies (9.43), substitute (9.42) into (9.36) to give

$$\begin{aligned} A_{conv}(t,\{s_m\}) &= t^{-1} \odot A(\{\bar{s}_m\}) \\ &= t^{-1} \odot A(\{\bar{f} \odot h_m + \bar{c}_m\}) \ . \end{aligned} \tag{9.45}$$

Assuming that (9.44) is valid for the effective contamination $\bar{c}_m$, and invoking (9.40) and (1.36), (9.45) becomes

$$\begin{aligned} A_{conv}(t,\{s_m\}) &= t^{-1} \odot \bar{f} \\ &= t^{-1} \odot (t \odot f) \\ &= (t^{-1} \odot t) \odot f \end{aligned}$$

$$= f \tag{9.46}$$

and (9.43) is seen to hold. The converse: that (9.43) implies (9.44), may be proved by retracing these steps.

9.4.3 Inverse Averaging

If convolution shift-and-add is to be used to reconstruct an image of an object without artefacts, the effective object must be a delta function. To achieve this, the function t should ideally be chosen as the convolutional inverse of f. Obviously, since it is f that is to be reconstructed, this choice cannot actually be made. However, if an estimate $\hat{f}$ of the object is available, one can define a reconstruction scheme, called 'inverse' averaging, based upon the choice

$$t = \hat{f}^{-1} , \tag{9.47}$$

which applies to any of the averaging techniques. The operator A_{inv} for inverse averaging is defined by

$$A_{inv}(\hat{f},\{s_m\}) = A_{conv}(\hat{f}^{-1},\{s_m\}) \tag{9.48}$$

and its effective object is

$$\bar{f} = \hat{f}^{-1} \odot f . \tag{9.49}$$

Note that if the estimate is a δ function then A_{inv} produces the same result as A,

$$A_{inv}(\delta,\{s_m\}) = A(\{s_m\}) . \tag{9.50}$$

9.4.4 Correlation Averaging

It is well known that inverse filtering tends to be appreciably noise-sensitive. Unless precautions are taken the filtered signal can be obliterated by the filtered noise. A similar situation could arise with inverse averaging. The effective contamination could dominate the effective object even though the actual contamination is small compared to the actual object.

An alternative to the choice (9.47) is

$$t = \hat{f}^{\dagger} , \tag{9.51}$$

where $\dagger$ denotes the adjoint operation defined by (1.51). The rationale behind this choice is made apparent by invoking (1.39) to express (9.34) as

$$T(g) = \hat{f}^{\dagger} \odot g$$

$$= \hat{f} * g \ . \tag{9.52}$$

The transformation selected by (9.51) is cross-correlation with the estimate $\hat{f}$. The algorithm described by (9.52) is called 'correlation' averaging and its operator A_{cor} is defined by

$$A_{cor}(\hat{f},\{s_m\}) = A_{conv}(\hat{f}^{\dagger},\{s_m\}) \ . \tag{9.53}$$

The effective object is

$$\bar{f} = \hat{f} * f \ . \tag{9.54}$$

A formula similar to (9.50) holds for A_{cor}:

$$A_{cor}(\delta,\{s_m\}) = A(\{s_m\}) \ . \tag{9.55}$$

The efficacy of correlation averaging can be seen by considering the case in which the estimate is equal to the object. In this case the object is the autocorrelation of the object. For some objects shift-and-add reconstructs f * f more faithfully than f (Chapter 7). The effective contamination is the cross-correlation of the object with the actual contamination. Provided the actual contamination is uncorrelated, and small compared to the actual object, the effective contamination is also small relative to the effective object.

9.4.5 Differential Averaging

In this subsection the choice

$$t_i = \frac{\partial \delta}{\partial x_i} \ , \qquad 1 \leq i \leq N \ , \tag{9.56}$$

where x_i is a coordinate, is considered. The convolution averaging operator constructed from an averaging operator A using the choice (9.56) is called the 'i^{th} differential' averaging operator, and is denoted by $A_{\partial i}$. In two dimensions, A_{∂_1} may be written $A_{\partial x}$ and A_{∂_2} as $A_{\partial y}$. In one dimension, A_{∂_1} is written A_{∂}.

The transformation corresponding to (9.56) is partial differentiation, as can be seen by substituting (9.56) into (9.34) and invoking (1.47), i.e.

$$T_i(g) = \frac{\partial \delta}{\partial x_i} \odot g$$

$$= \delta \odot \frac{\partial g}{\partial x_i}$$

$$= \frac{\partial g}{\partial x_i} \quad . \tag{9.57}$$

So, the inverse transformation is integration with respect to the i^{th} coordinate, i.e.

$$T_i^{-1}(g)(\mathbf{r}) = \int_{-\infty}^{x_i} dx_i' \; g(\mathbf{r}') \tag{9.58}$$

where

$$\mathbf{r} = (x_1, \ldots, x_N) \tag{9.59}$$

and

$$\mathbf{r}' = (x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_N) \quad . \tag{9.60}$$

The convolutional inverse of the filter (9.56) is

$$t^{-1}(\mathbf{r}) = \int_{-\infty}^{x_i} dx_i' \; \delta(\mathbf{r}') \tag{9.61}$$

$$= \text{step}(x_i) \prod_{k=1, k\neq i}^{N} \delta(x_i) \quad . \tag{9.62}$$

To verify this, convolve t_i and t_i^{-1} and invoke (1.47),

$$(t_i \odot t_i^{-1})(\mathbf{r}) = \frac{\partial \delta}{\partial x_i} \odot \int_{-\infty}^{x_i} dx_i' \; \delta(\mathbf{r}')$$

$$= \delta \odot \frac{\partial}{\partial x_i} \int_{-\infty}^{x_i} dx_i' \; \delta(\mathbf{r}')$$

$$= \delta \odot \delta$$

$$= \delta \quad . \tag{9.63}$$

A formula for $A_{\partial i}$ is obtained by substituting (9.57) and (9.58) into (9.33), i.e.

$$A_{\partial i}(\{s_m\})(\mathbf{r}) = \int_{-\infty}^{x_i} dx_i' \, A(\{\frac{\partial s_m}{\partial x_i}\})(\mathbf{r}') \ . \tag{9.64}$$

From (9.64) it is apparent that differential averaging consists of differentiating the speckle images with respect to x_i, averaging, and then integrating with respect to x_i. The adjective 'differential' in the term 'differential averaging' is justified by this sequence of operations.

The rationale for differential averaging may be summed up by a simple theorem. If the partial derivative of an object f with respect to x_i can be reconstructed by an averaging operator A, i.e.

$$\frac{\partial f}{\partial x_i} \text{ is a member of } R(A) \ , \tag{9.65}$$

then the object can be reconstructed by differential averaging, i.e.

$$f \text{ is a member of } R(A_{\partial i}) \ . \tag{9.66}$$

The theorem may be proved by considering the partial derivative of each speckle image of f. The result of applying (1.47) to (9.1) is

$$\begin{aligned} \frac{\partial s_m}{\partial x_i} &= \frac{\partial}{\partial x_i}(f \odot h_m + c_m) \\ &= \frac{\partial f}{\partial x_i} \odot h_m + \frac{\partial c_m}{\partial x_i} \ . \end{aligned} \tag{9.67}$$

Since $\partial f/\partial x_i$ can be reconstructed by A, one has

$$\frac{\partial f}{\partial x_i} = A(\{\frac{\partial f}{\partial x_i} \odot h_m + d_m\}) \ , \tag{9.68}$$

where d_m is the m^{th} contamination. The assumption that (9.68) is valid for the contamination d given by

$$d_m = \frac{\partial c_m}{\partial x_i} \tag{9.69}$$

is now made. Before the theorem can be invoked, the validity of this assumption needs to established for the particular contaminations present in the situation of interest. However, in general it seems reasonable that the assumption holds if the actual contaminations c_m are sufficiently small. Certainly, it holds if all of the contaminations vanish.

Substituting (9.69) into (9.68) and invoking (9.67), the relationship

$$\frac{\partial f}{\partial x_i} = A\left(\left\{\frac{\partial f}{\partial x_i} \odot h_m + \frac{\partial c_m}{\partial x_i}\right\}\right)$$

$$= A\left(\left\{\frac{\partial s_m}{\partial x_i}\right\}\right) \tag{9.70}$$

is found. Finally, (9.70) is substituted into (9.64) to determine the differential average of the ensemble $\{s_m\}$, i.e.

$$A_{\partial i}(\{s_m\})(\mathbf{r}) = \int_{-\infty}^{x_i} dx_i' \frac{\partial f_m}{\partial x_i}(\mathbf{r}')$$

$$= f(\mathbf{r}) \ . \tag{9.71}$$

Hence, the object f may be reconstructed, and the theorem is proved. The theorem may also be established from the theorem of ¶9.4.2 by noting from (9.57) and (9.40) that the effective object for differential averaging is

$$\bar{f} = \frac{\partial f}{\partial x_i} \ . \tag{9.72}$$

The theorem has been derived from first principles because a proof based on the derivatives of delta functions is apt to be somewhat unconvincing.

9.5 AN ITERATIVE SCHEME

The correlation and inverse averaging techniques require an estimate $\hat{f}$ of the object. Since a new estimate of the object is produced by these techniques, an iterative scheme can be defined. The first estimate $\hat{f}^{(1)}$ is obtained by averaging the ensemble with A, and applying the extent image operator E (¶7.4.2), i.e.

$$\hat{f}^{(1)} = E[A(\{s_m\})] \ . \tag{9.73}$$

The second and third arguments of E as defined by (7.34) are omitted for brevity. Subsequent estimates are calculated from

$$\hat{f}^{(i)} = E[A_{ave}(\hat{f}^{(i-1)}, \{s_m\})] \ , \quad i \geq 2 \ , \tag{9.74}$$

where the i^{th} estimate is written $\hat{f}^{(i)}$ and A_{ave} denotes either A_{cor} or A_{inv}.

A more compact description of the iterative scheme is possible. To achieve it, note that (9.50) and (9.55) allow A to be expressed in terms of A_{ave} as

$$A(\{s_m\}) = A_{ave}(\delta, \{s_m\}) . \tag{9.75}$$

If the zeroth estimate is defined to be a δ function, (9.73) may be written in the form of (9.74). An alternative description of the iterative scheme is therefore

$$\hat{f}^{(0)} = \delta \tag{9.76}$$

and

$$\hat{f}^{(i)} = E[A_{ave}(\hat{f}^{(i-1)}, \{s_m\})] , \quad i \geq 1 . \tag{9.77}$$

An advantage of the description implied by (9.76) and (9.77) over (9.73) and (9.74) is that a priori assumptions about the nature of the object are more apparent. For instance, (9.76) can be interpreted as the assumption that the object contains at least one δ function.

The iterative scheme is illustrated in Fig. 9.3, using the conventions established in ¶9.3. The transformation T indicated in the figure is defined by

$$T(g) = \hat{f} * g \tag{9.78}$$

for correlation averaging or

$$T(g) = \hat{f}^{-1} \odot g \tag{9.79}$$

for inverse averaging.

The iterative scheme may be represented by a single averaging operator $A_{ave}^{(i)}(\hat{f}^{(0)}, \cdot)$, which maps $\{s_m\}$ to $\hat{f}^{(i)}$ given an initial estimate $\hat{f}^{(0)}$, i.e.

$$\hat{f}^{(i)} = A_{ave}^{(i)}(\hat{f}^{(0)}, \{s_m\}) . \tag{9.80}$$

If the sequence $\{\hat{f}^{(i)}\}$ converges to an image $\hat{f}$, an operator $A_{ave}^{(\infty)}(\hat{f}^{(0)}, \cdot)$ may be defined through

$$\hat{f} = A_{ave}^{(\infty)}(\hat{f}^{(0)}, \{s_m\}) . \tag{9.81}$$

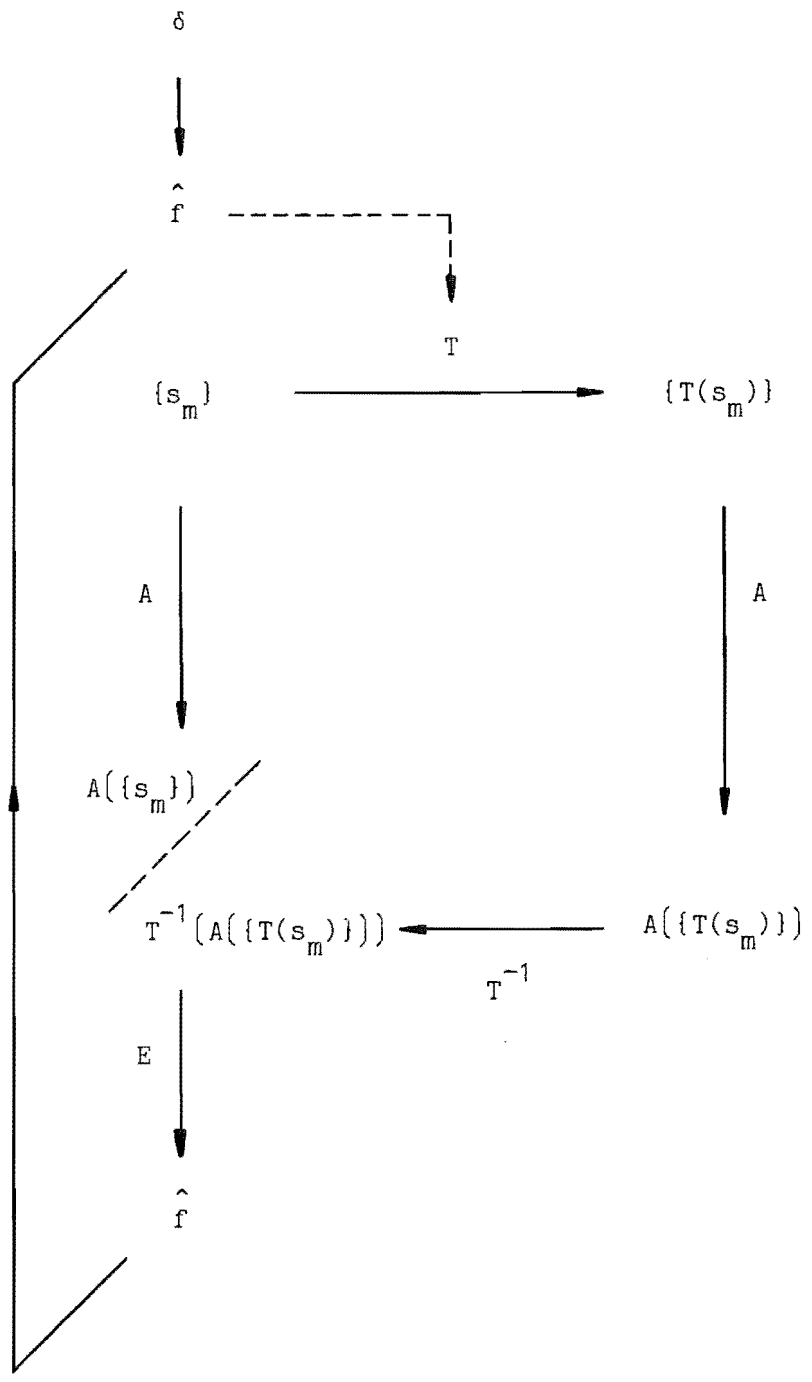


Fig. 9.3 The iterative transformation technique. The transformation T requires an estimate of the object. This is indicated by the dashed arrow. The initial estimate is a δ function. A new estimate is obtained from the transform average by applying an extent constraint E.

9.6 COMPOSITE AVERAGING

Two convolution averaging operators may be combined to form another. Let the two filter functions be t_1 and t_2. The combined operator is denoted by $A_{conv,conv}(t_1,t_2)$ and is defined by the formula

$$A_{conv,conv}(t_1,t_2,\{s_m\}) = A_{conv}(t_1 \circledcirc t_2,\{s_m\}) . \tag{9.82}$$

A natural extension of the notation $A_{conv,conv}$ allows a filter function to be omitted if the corresponding convolution average is a differential average. For example,

$$A_{\partial i,conv}(t,\{s_m\}) = A_{conv,conv}(\frac{\partial \delta}{\partial x_i},t,\{s_m\}) \tag{9.83}$$

and

$$A_{conv,\partial i}(t,\{s_m\}) = A_{conv,conv}(t,\frac{\partial \delta}{\partial x_i},\{s_m\}) . \tag{9.84}$$

An additional extension of the notation $A_{conv,conv}$ allows correlation (¶9.4.4) and inverse (¶9.4.3) averaging to be specified as one or both convolution averages. For example,

$$A_{conv,cor}(t,\hat{f},\{s_m\}) = A_{conv,conv}(t,\hat{f}^{\dagger},\{s_m\}) \tag{9.85}$$

and

$$A_{conv,inv}(t,\hat{f},\{s_m\}) = A_{conv,conv}(t,\hat{f}^{-1},\{s_m\}) . \tag{9.86}$$

9.7 AN ITERATIVE PROCEDURE BASED UPON COMPOSITE AVERAGING

This section formalises the following suggestion.

Suppose it is found that the iterative scheme of ¶9.5 is capable of reconstructing a certain class C_1 of objects using an averaging operator A_{ave}. Suppose also that it is noted the a transformation S transforms an object of another class C_2 into an object belonging to C_1. The suggestion is that, by applying the same transformation to an ensemble of speckle images of an object in C_2, the object may perhaps be reconstructed.

The algorithm based upon this suggestion is depicted in Fig. 9.4. The transformation T shown in the figure is defined by (9.78) or (9.79),

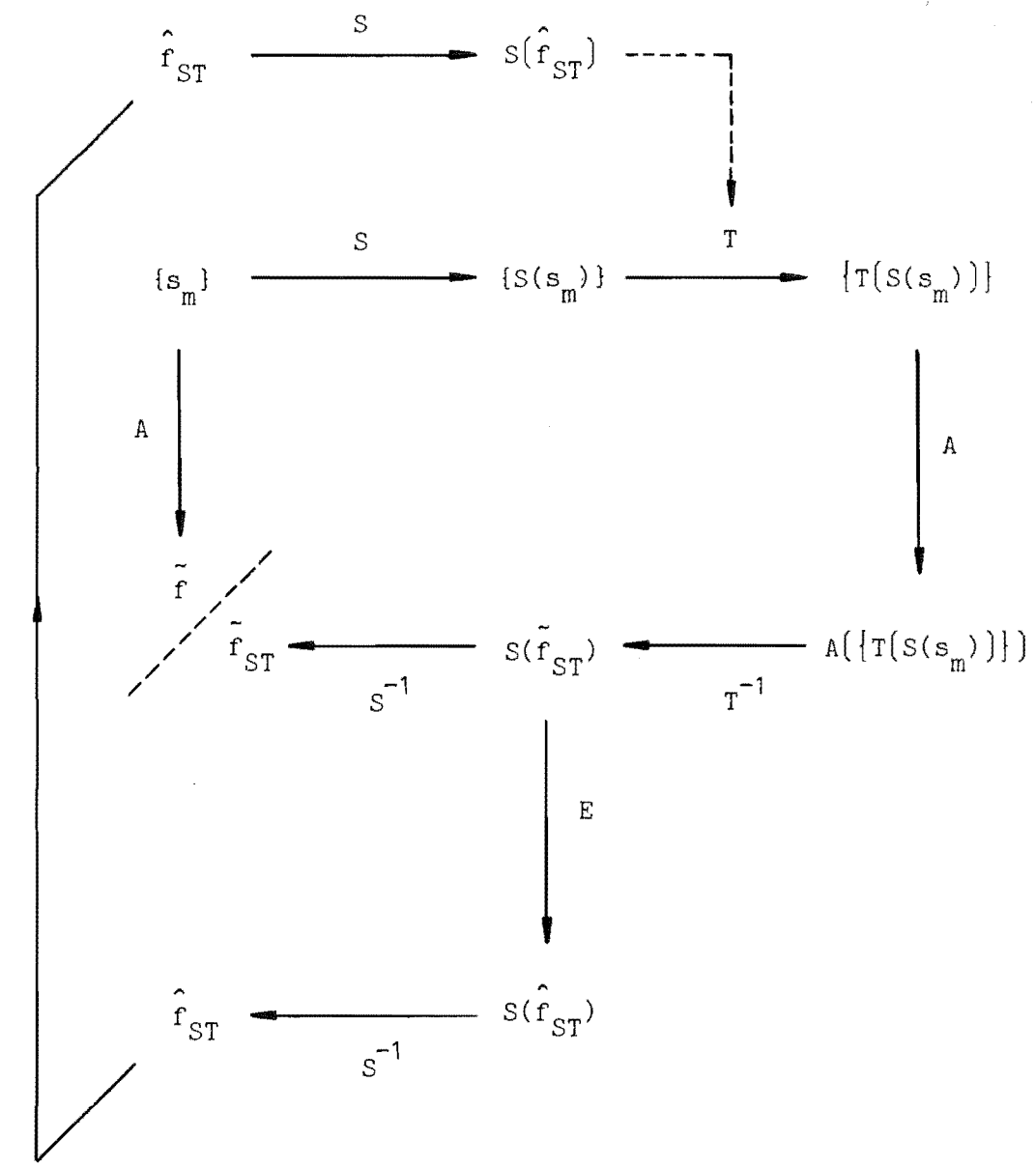


Fig. 9.4 An iterative scheme based on composite averaging. The composite average is formed from two transformations S and T, and an averaging operator A.

depending on the nature of A_{ave}. For a transformation S using a filter function s, the algorithm is written as

$$\hat{f}^{(0)} = s^{-1} \tag{9.87}$$

and

$$\hat{f}^{(i)} = s^{-1} \odot E[s \odot A_{conv,ave}(s, s \odot \hat{f}^{(i-1)}, \{s_m\})] \ . \tag{9.88}$$

The composite operator $A_{conv,ave}$ combines the transformation s with the transformation, (9.78) or (9.79), implicit in the A_{ave} operator. The operator $A^{(i)}_{conv,ave}(s, \hat{f}^{(0)}, \cdot)$ maps the ensemble to $\hat{f}^{(i)}$ given the filter function s and initial estimate $\hat{f}^{(0)}$, i.e.

$$\hat{f}^{(i)} = A^{(i)}_{conv,ave}(s, \hat{f}^{(0)}, \{s_m\}) \ . \tag{9.89}$$

When the sequence $\{\hat{f}^{(i)}\}$ converges to an image $\hat{f}$, an operator $A^{(\infty)}_{conv,ave}(s, \hat{f}^{(0)}, \cdot)$ may be defined such that

$$\hat{f} = A^{(\infty)}_{conv,ave}(s, \hat{f}^{(0)}, \{s_m\}) \ . \tag{9.90}$$

The definitions (9.89) and (9.90) are analogous to (9.80) and (9.81) for the iterative scheme described in ¶9.5.

The suggestion made at the beginning of the section is in fact correct and it is encapsulated in the following theorem. An object f can be reconstructed from an ensemble $\{s_m\}$ by the scheme (9.88), with initial estimate $\hat{f}^{(0)}$, i.e.

$$f \text{ is a member of } R(A^{(\infty)}_{conv,ave}(s, \hat{f}^{(0)}, \cdot)) \ , \tag{9.91}$$

if $s \odot f$ can be reconstructed from $\{s \odot s_m\}$ by the scheme using $s \odot \hat{f}^{(0)}$ as an initial estimate, i.e.

$$s \odot f \text{ is a member of } R(A^{(\infty)}_{ave}(s \odot \hat{f}^{(0)}, \{s \odot s_m\})) \ . \tag{9.92}$$

To prove this, first consider the effect of A_{conv} on an ensemble $\{s \odot s_m\}$. Two invocations of the definition of A_{conv} (9.36) lead to

$$\begin{aligned} A_{conv}(t, \{s \odot s_m\}) &= t^{-1} \odot A(\{t \odot (s \odot s_m)\}) \\ &= s \odot (s \odot t)^{-1} \odot A(\{(s \odot t) \odot s_m\}) \\ &= s \odot A_{conv}(s \odot t, \{s_m\}) \ . \end{aligned} \tag{9.93}$$

Define g, $\hat{g}^{(i)}$ and $\{t_m\}$ by

$$g = s \odot g \ , \tag{9.94}$$

$$\hat{g}^{(i)} = s \odot \hat{f}^{(i)} \ , \tag{9.95}$$

and

$$t_m = s \odot s_m \ . \tag{9.96}$$

When both sides of (9.88) are convolved with s and (9.95) is invoked, the result

$$\hat{g}^{(i)} = E[s \odot A_{conv,ave}(s, \hat{g}^{(i-1)}, \{s_m\})] \tag{9.97}$$

is obtained. Let the filter function for A_{ave} be denoted by t. For instance, for $A_{ave} = A_{inv}$,

$$t = \hat{g}^{(i-1)-1} \ , \tag{9.98}$$

and for $A_{ave} = A_{cor}$,

$$t = \hat{g}^{(i-1)\dagger} \ . \tag{9.99}$$

The operand of the extent image operator in (9.97) can be expressed in terms of $A_{conv,conv}$ with the aid of (9.85) or (9.86). The definition (9.82) of $A_{conv,conv}$ is substituted into the result and (9.93) is invoked to produce

$$\begin{aligned} s \odot A_{conv,ave}(s, \hat{g}^{(i-1)}, \{s_m\}) &= s \odot A_{conv,conv}(s, t, \{s_m\}) \\ &= s \odot A_{conv}(s \odot t, \{s_m\}) \\ &= A_{conv}(t, \{s \odot s_m\}) \\ &= A_{ave}(\hat{g}^{(i-1)}, \{s \odot s_m\}) \ . \end{aligned} \tag{9.100}$$

After invoking the definition (9.96) of $\{t_m\}$ and substituting (9.100) into (9.97) one obtains

$$\hat{g}^{(i)} = E[A_{ave}(\hat{g}^{(i-1)}, \{t_m\})] \tag{9.101}$$

or

$$\hat{g}^{(i)} = A_{ave}^{(i)}(\hat{g}^{(0)}, \{t_m\}) \tag{9.102}$$

where (9.80) invoked. Because of (9.92), $\hat{g}^{(i)}$ converges to g,

$$\hat{g}^{(i)} \to g \ , \tag{9.103}$$

and therefore

$$\hat{f}^{(i)} \to f \ . \tag{9.104}$$

(9.104) implies that f is reconstructed by $A^{(\infty)}_{conv,ave}(s,\hat{f}^{(0)},\cdot)$ and that (9.91) holds, thus proving the theorem.

By substituting (9.100) into (9.88), an alternative formulation of the iterative scheme is obtained,

$$\hat{f}^{(0)} = s^{-1} \tag{9.105}$$

$$\hat{f}^{(i)} = s^{-1} \odot E[A_{ave}(s \odot \hat{f}^{(i-1)}, \{s \odot s_m\})] \ . \tag{9.106}$$

The significance of this formulation is that it shows that $\hat{f}^{(i)}$ may be calculated using A_{ave} instead of $A_{conv,ave}$.

9.8 RECONSTRUCTIBLE OBJECTS

The simulation results reported in ¶8.2 show that a significant proportion of one-dimensional objects of the form

$$f(\mathbf{r}) = \sum_i a_i \delta(\mathbf{r} - \mathbf{r}_i) \tag{9.107}$$

may be reconstructed by the iterative scheme of ¶9.5. The initial estimate is a δ function and inverse averaging is used. This may be written,

$$f \text{ is a member of } R(A^{(\infty)}_{inv}(\delta,\cdot)) \tag{9.108}$$

using the notation of ¶9.5. The object defined by (9.107) may also be written

$$f = \sum_i a_i T_{\mathbf{r}_i}(\delta) \tag{9.109}$$

with the aid of (1.31).

Consider a one-dimensional object g

$$g = \sum_i a_i T_{\mathbf{r}_i}(\text{step}) \tag{9.110}$$

which is piecewise constant. The derivative g' of g is an array of δ functions of the form (9.109). Hence, g may be reconstructed by the extended iterative scheme of ¶9.7 with the transformation

$$s = \delta' \tag{9.111}$$

which corresponds to differentiation. The initial estimate is a step function and inverse averaging is used. This is written

$$g \text{ is a member of } R\left[A^{(\infty)}_{\partial,\text{inv}}(\text{step},\cdot)\right] \tag{9.112}$$

using the abbreviation A_∂ for $A_{\partial i}$ (¶9.4.5) and omitting δ' from the denotation of the operator in the manner indicated by (9.83). The iterative step for this operator is obtained by substituting (9.111) into (9.106),

$$\hat{f}^{(i)} = \int E\left[A_{\text{inv}}\left[\hat{f}^{(i-1)}, \{s'_m\}\right]\right] . \tag{9.113}$$

This is a concise formulation of the differential technique investigated in ¶8.3 for which the averaging operator A is SAA.

9.9 CONCLUSIONS

In many situations an ensemble of independently distorted images of an object can be obtained. If, for each image, the phase of the Fourier transform of the distortion is contaminated with a significant random component then the distortion is called a 'speckle' point-spread function (PSF) and the images are called 'speckle' images. The reconstruction of the object from the speckle images is called the 'speckle imaging' problem.

A theoretical solution to the speckle imaging problem called 'ideal' shift-and-add is presented in ¶6.4.3. The solution requires that the location of the brightest point of each speckle PSF be known. In practice, this information is not available. A method called 'weighted' shift-and-add is discussed in ¶6.4.4 which is an approximate implementation of ideal shift-and-add. The approximation causes the images of the objects to contain a ghosting artefact. It is shown in ¶6.8 that, in the absence of ghosting, weighted shift-and-add in conjunction with Fourier imaging (discussed in ¶1.7) is an approximate solution to the remote probing problem (¶1.5).

In Chapters 6, 7, and 8 the ghosting artefact is investigated in detail. A theoretical model to account for it is presented in ¶6.7 and is validated by simulation. In Chapters 7 and 8 extensions of weighted shift-and-add are presented for reducing the ghosting artefact. A method called 'correlation' shift-and-add (Chapter 6) allows a significant proportion of objects dominated by several isolated points of arbitrary brightness to be reconstructed. However, the amplitudes of the points are not given accurately by the method. A second method, called 'inverse'

shift-and-add (¶8.2), theoretically allows arbitrary objects to be reconstructed. In practice, useful reconstructions of objects dominated by several isolated bright points are obtained. A differential technique (¶8.3) allows a significant fraction of the class of piecewise constant objects to be reconstructed.

A theoretical basis for the new techniques is established in this chapter. They are shown to be equivalent to performing ordinary shift-and-add on a transformed ensemble of speckle images and then applying the inverse transformation to the result of shift-and-add. The differential technique (¶8.3) comprises two such transformations.

The support for the conclusions expressed here consists of extensive simulations in one-dimension, experimental results in two dimensions, and theoretical analysis. The experimental results show that naturally occurring biological structures can serve as a reference for shift-and-add and that correlation shift-and-add can be usefully applied to experimental data.

10. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

10.1 INTRODUCTION

The research conducted for this thesis has, as is usual, raised as many questions as it has answered. The differential technique (¶8.3) has yet to be tested on actual data. The difficulty here lies in the construction of suitable phantoms. Some experience in this direction is described in ¶10.2. The correlation technique (Chapter 7) exhibits a stable monotonic convergence. Methods for accelerating the convergence are proposed in ¶10.3. ¶10.4 describes a refinement of the extent constraint (¶7.4.1) which is used to obtain an estimate from the shift-and-add image in the generalised iterative transform technique (Chapter 9). The refraction correction introduced ¶8.4 needs to be tested on actual data. Some ideas for the construction of suitable phantoms are presented in ¶10.5. Applications, other than in ultrasonic imaging, for the techniques investigated in this thesis are considered in ¶10.6. Final conclusions are presented in ¶10.7.

10.2 CONSTRUCTION OF EXTENDED ULTRASONIC OBJECTS

Most of the experimental images presented in this thesis are obtained from ultrasonic objects consisting of metal strips and wires. The cross-sections of these objects are lines and points respectively. To obtain an extended object, a three-dimensional target which scatters significantly throughout its volume is required. A target consisting of a cloud of closely spaced scatterers approximately satisfies this requirement. The author has investigated a method for constructing such targets. The method and the results obtained are the subject of this section.

The scatterers chosen for the target are silicon carbide (SiC) particles. These were chosen because they are readily available in a wide range of grit sizes. Three grit sizes were given particular attention because they happen to be of the same order as the minimum wavelength λ_{min} of the experimental imaging system described in Chapter 5. Different grit sizes are graded by the manufacturers (Carborundum, Manchester) by using sieves of varying mesh sizes. The grit size is specified by a standard code number which relates to the size of the mesh. The sizes used in the targets described in this section are '60', '100',

and '220'. The average particle sizes for the selected grit sizes are (Coes 1971 p.62) 406 μ, 173 μ and 66 μ respectively, where μ = 0.001 mm. In terms of λ_{min}, these sizes are 0.92 λ_{min}, 0.39 λ_{min} and 0.15 λ_{min}, respectively. Hence, the 220-grit particles are effectively point scatterers while the 60-grit particles may scatter anisotropically. Particles of SiC are multi-faceted rather than smooth. For 60-grit particles the scattering patterns could therefore vary appreciably with angle.

The SiC particles are suspended in gelatine. This substance was chosen because it has a similar acoustic impedance to water and may be melted and solidified at conveniently low temperatures. Two concentrations, c_c and c_s, of SiC particles are used. For the concentration c_c, 100 ml of the suspension contains 1.2 g of SiC, while for c_s the same volume of suspension contains 41 g of SiC. The rationale for this choice is discussed below.

The suspensions are placed in small plastic boxes which are designed to be supported by the cassette method described in ¶5.2.3. In the x-y plane (refer to Fig. 5.2) the cross-section of each box is 3 mm × 5 mm. The length of the box is 48 mm. A single piece of 0.2 mm thick mylar is folded and glued with araldite to form each box.

The scattering from the filled box is expected to arise from the SiC particles suspended in the gelatine and from the walls of the box. The chosen concentrations of SiC particles were estimated to be such as to make the two contributions of comparable magnitude. The corners of the box are expected to scatter significant intensities of ultrasound in most directions in the x-y plane (invoking the concepts of the geometrical theory of diffraction). Since the walls of the box are of a finite width, the 'corners' of the box can be considered to occupy a finite volume (see Fig. 10.1). The volume of the 'corners' is calculated by assuming that they have a 0.2 mm × 0.2 mm square cross-section and are 48 mm long. The concentration c_c of SiC particles was chosen so that the volume of SiC particles within the box was approximately equal to that of the 'corners'. The concentration c_s of SiC particles was chosen to correspond, in similar fashion, to the volume of the 'sides' of the box, The latter volume was calculated on a similar basis to that invoked for the 'corners'.

Suspending the SiC particles in the gelatine is a difficult procedure. The SiC particles and gelatine must be mixed while the gelatine is in a viscous state, but this introduces air bubbles. Most of these air

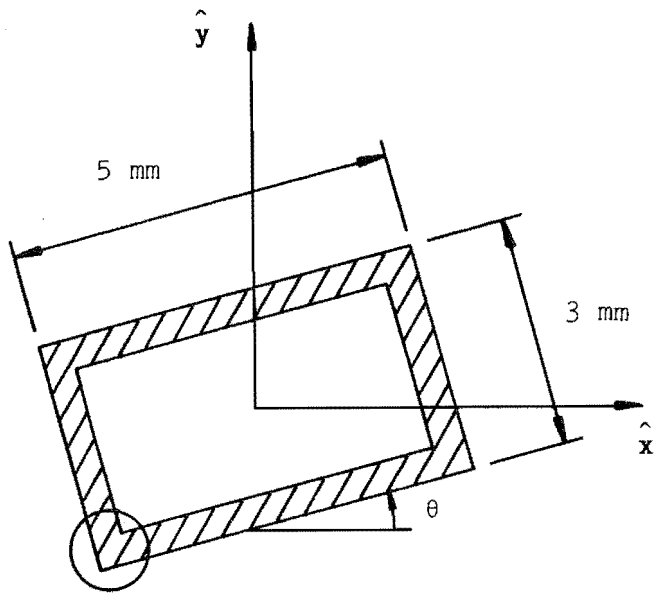


Fig. 10.1 A cross-section through the box target in the x-y plane. The circle indicates the location of the glued edge. The walls (shaded) of the box are made of 0.2 mm thick mylar.

bubbles can be removed by placing the suspension in a vacuum. This is achieved by placing the suspension (and the container in which it is mixed) into a vessel connected to a vacuum pump. The pump lowers the pressure in the vessel to $< 10^{-2}$ mm Hg. During this 'degassing' process the air bubbles in the suspension expand. Eventually, the suspension bursts open to allow the air within the bubble to escape and be removed by the pump. The suspension must be placed in a large container for the degassing to be effective, so the suspension cannot be degassed while inside the target box. Additionally, the suspension is distributed over the bottom and sides of the container by the degassing procedure. After degassing, the suspension is solid, but is returned to the viscous state by placing the container in a warm (≈35°C) water bath. The viscous suspension is injected into the box with a syringe. The box is then placed in cold running water to solidify the contents. Finally, the ends of the box are closed and glued. The box must be airtight because otherwise the gelatine gradually looses its moisture content and shrinks.

There is an alternative procedure which is begun by placing the appropriate quantities of SiC particles and liquid gelatine into the box. The box and contents are then cooled and the ends glued shut. Next, the gelatine within the box is melted. By judicious rotation of the box the SiC particles can be distributed throughout the gelatine. While the box is still being rotated it is cooled, so as to 'freeze' the suspension. The drawback of this method is that the rotation of the box does not necessarily guarantee that the distribution of the scatterers is random. If in a particular instance, however, the resulting distribution does not appear to be random, the suspension may be liquified and the process repeated. Eventually, an apparently random distribution may be achieved.

The orientation of the box is specified as shown in Fig. 10.1. Each target was scanned for $\theta = 0°$ and additionally a box filled with water was scanned at $\theta = 0°$, $45°$, $-45°$ and $-90°$. The measured data were processed by elementary shift-and-add and the resulting images displayed on a video monitor. In all cases, the images consisted of one, two or three peaks, which may well be due to scattering from the corners of the box. The images did not contain the extended image of the SiC particles that was expected. The most probable explanation for this is that the scattering from the particles was considerably weaker than that from the box. The author suggests that further experiments be carried out with higher concentrations of SiC particles. Additionally, a wire 'reference' could be introduced into the object region (¶5.2.3) to assist in the identification of the corners of the box.

10.3 CONVERGENCE ACCELERATION

The analyses presented for iterative correlation (¶7.5) and inverse (¶8.2.4) shift-and-add for two-point objects show that the estimate $\hat{f}^{(i)}$ and the shift-and-add image $\tilde{f}^{(i)}$ converge linearly, i.e. the error in the $(i+1)^{th}$ iteration is proportional to the error in the i^{th} iteration. There are standard methods for accelerating the convergence of such sequences (cf. Conte and de Boor 1972). These methods were not investigated because it was felt that it was more important to determine the types of objects which can be successfully reconstructed. However, in an application where speckle images are to be processed by an iterative shift-and-add technique on a routine basis, a method to accelerate the convergence would be useful. It is shown in this section how a standard convergence acceleration method could be applied to any one of the

iterative shift-and-add techniques investigated in this thesis.

One of the standard methods for accelerating the convergence of a sequence which is converging linearly is called 'Steffenson iteration' (Conte and de Boor 1972 p.54). It is based on Aitken's Δ^2 process (Conte and de Boor 1972 p.53) and could be applied to either of the sequences $\{\hat{f}^{(i)}\}$ and $\{\tilde{f}^{(i)}\}$. For illustration, it is applied to $\{\hat{f}^{(i)}\}$ here. Define $\Delta\hat{f}^{(i)}$ by

$$\Delta\hat{f}^{(i)} = \hat{f}^{(i+1)} - \hat{f}^{(i)} . \tag{10.1}$$

Starting with the estimate $\hat{f}^{(i-2)}$, two iterations of correlation shift-and-add (for example) are performed. This yields the estimates $\hat{f}^{(i-1)}$ and $\hat{f}^{(i)}$. The estimate $\hat{f}^{(i+1)}$ is then calculated from

$$d = \Delta\hat{f}^{(i-1)} \tag{10.2}$$

$$r = \Delta\hat{f}^{(i-2)}/d \tag{10.3}$$

and

$$\hat{f}^{(i+1)} = \hat{f}^{(i)} + d/(r-1) . \tag{10.4}$$

The process is then repeated, i.e. two iterations of correlation shift-and-add are calculated with $\hat{f}^{(i+1)}$ as initial estimate, and so on.

The method is expected to be helpful for the two-point objects on which the analyses in ¶¶7.5 and 8.2.4 are based. The convergence of iterative correlation and inverse shift-and-add for more general objects is, in the author's experience, often approximately linear. Therefore, the modified iterative method defined by (10.1) through (10.4) is expected to be generally useful.

10.4 AN IMPROVEMENT ON THE EXTENT CONSTRAINT

An essential step in the generalised shift-and-add algorithm described in this thesis is the extraction of an estimate $\hat{f}^{(i)}$ of the object from a shift-and-add image $\tilde{f}^{(i)}$ obtained using an estimate $\hat{f}^{(i-1)}$. $\hat{f}^{(i)}$ is determined by an extent constraint as explained in ¶7.4.1. In this section, a more sophisticated method for obtaining $\hat{f}^{(i)}$ is proposed. The author did not have time to investigate it, but believes it to be particularly promising.

The case of correlation shift-and-add is considered for illustrative purposes. When (7.21) is invoked, the approximate model (7.22)

may be written as

$$\tilde{f}^{(i)} = [1 - \tfrac{1}{2}\int|\bar{g}^{(i)}|]f + \tfrac{1}{2}|g^{-(i)}| * f . \qquad (10.5)$$

Note that as the effective object $\bar{f}(i)$ tends to a δ function, the shift-and-add image $\tilde{f}^{(i)}$ tends to the object f. Ideally, one would like to solve (10.5) for f. This is not directly possible because $\bar{f}^{(i)}$ and $\bar{g}^{(i)}$ depend on f via (7.6). However, an iterative procedure may be used to obtain an approximate solution. An initial estimate $\hat{f}_0^{(i)}$ of the object is obtained by applying the extent constraint, i.e.

$$\hat{f}_0^{(i)} = E\{\tilde{f}^{(i)}\} \qquad (10.6)$$

where the extent operator E is defined in ¶7.4.2. A sequence $\{\hat{f}_k^{(i)}\}$ for k = 0, 1, ... of estimates of the object are then calculated as follows. Using the estimate $\hat{f}_k^{(i)}$, a further estimate $\hat{\bar{f}}_k^{(i)}$ of the effective object is given by (cf. (7.6))

$$\hat{\bar{f}}_k^{(i)} = \hat{f}_k^{(i-1)} * f . \qquad (10.7)$$

An image $\hat{\bar{g}}_k^{(i)}$ is computed according to

$$\hat{\bar{g}}^{(i)} = \hat{\bar{f}}^{(i)} - \delta . \qquad (10.8)$$

A potentially better estimate of the object is then obtained from

$$\hat{f}^{(i)} = \frac{\tilde{f}^{(i)} - \tfrac{1}{2}|\hat{\bar{g}}_k^{(i)}| * f}{1 - \tfrac{1}{2}\int|\hat{\bar{g}}_k^{(i)}|} . \qquad (10.9)$$

Since the model (10.5) is only approximate, probably only $\hat{f}_1^{(i)}$ or $\hat{f}_2^{(i)}$ need be calculated. When $\bar{g}^{(i)}$ is small, (10.5) is expected to be a more accurate model. In this case more members of the sequence $\{\hat{f}_k^{(i)}\}$ may profitably be calculated. If $\hat{f}_K^{(i)}$ is the last member of the sequence calculated, then $\hat{f}^{(i)}$ is given by

$$\hat{f}^{(i)} = \hat{f}_K^{(i)} \qquad (10.10)$$

and another iteration of correlation shift-and-add is performed using $\hat{f}^{(i)}$ as the estimate.

The reason why this scheme may be better than just a simple application of the extent constraint can be seen from (10.9). In general, the second term in the numerator of the right hand side is nonzero inside

the region T_{obj} where the object is predicted to be by the extent constraint. Thus (10.9) allows ghosts which are in T_{obj} to be approximately removed. This is in contrast to the extent constraint which only removes ghosts from the region outside T_{obj}. The ghosts cannot be removed completely, however, because (10.5) is approximate. Hence, an iteration of correlation shift-and-add is required to obtain a new effective object $\bar{f}^{(i+1)}$, which is more like a δ function than $\bar{f}^{(i)}$. Then (10.5) can be expected to hold more accurately for the $(i+1)^{th}$ iteration than for the i^{th}. Thus the number K may be increased.

The iterative procedure described in this section should improve the convergence of iterative correlation and inverse shift-and-add because better estimates can be expected to be obtained from the shift-and-add images.

10.5 CONSTRUCTION OF PHANTOMS WITH A VARYING REFRACTIVE INDEX

The refractive index correction proposed in ¶8.4 needs to be tested on experimental data. A possible design for a suitable phantom is presented in this section.

The phantom might consist of two semicircular mylar buckets mounted as shown in Fig. 10.2. By filling the buckets with various liquids, two regions of different refractive indices could be obtained. Wires should be mounted in each bucket to act as a target. The incident beam may be reflected by the interface between the two buckets. However, the reflected beam would not enter the measurement aperture so this is not expected to cause any difficulty. Also, the multiple scattering caused by the reflected beam striking the target is expected to be rejected by shift-and-add.

10.6 APPLICATIONS OTHER THAN ULTRASONICS

Several extensions of shift-and-add are proposed and investigated in this thesis. The experimental data used to illustrate the various imaging techniques were obtained from ultrasonic scattering measurements. It is essential to bear in mind that the processing is coherent, in that the magnitude and phase of the speckle images are available. Since speckle processing is of at least as much interest in optical astronomy as in ultrasonics, it is important that the methods developed in this thesis be applied to optical data. Since optical images are incoherent, the contamination term c_{id} (defined by (6.23)) for ideal shift-and-add is

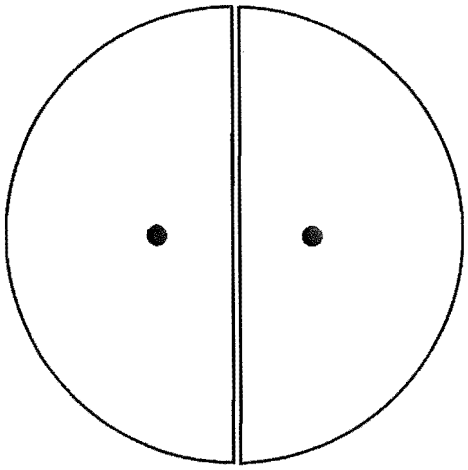


Fig. 10.2 A phantom with a varying refractive index. Each semicircle represents a mylar bucket filled with a liquid of a known refractive index. The target (shown by the dots) is mounted within the buckets. The beam is incident from the left.

nonzero. In weighted shift-and-add images the ghosting artefact and the object are superposed with a smooth background fog. When employing the iterative methods developed in this thesis, one would probably be best advised to attempt to reduce the amplitude of the fog by suitable filtering of the shift-and-add image before the estimate is obtained from (7.34).

Shift-and-add could well have applications in sciences other than ultrasonics and astronomy. For example, Szuszczewicz et al. (1983) consider the impact of ionospheric irregularities on synthetic aperture radars. They suggest that the irregularities could cause phase variations. These variations would distort the images obtained from such radars. The distortion is likely to be speckle-like. If this is the case, the methods proposed in this thesis could well enable the distorted images to restored.

10.7 CONCLUSIONS

A computer-aided imaging (CAI) system allows the nature of an object to be determined by measurements made some distance from it. To construct a CAI system a solution to the remote probing problem (¶1.5) is required. The solution is implemented as software on a computer and the computer is interfaced to a facility for acquiring the appropriate data.

The work reported in this thesis covers all of the above aspects. An idealised form of the shift-and-add principle (in conjunction with Fourier imaging) is shown to be an approximate solution of the remote probing problem. A model is derived which allows the deficiencies of ordinary shift-and-add to be analysed. A generalisation of shift-and-add is formulated which remedies some of these deficiencies. A software module has been written which facilitates the implementation of CAI algorithms. An ultrasonic measurement facility has been interfaced to a computer to allow the combination of computer and measurement facility to function as a CAI system.

The most significant aspect of the work reported the thesis is the generalised shift-and-add principle. Shift-and-add is an important speckle processing technique because it is more insensitive to contamination than other techniques such as Labeyrie processing (Labeyrie 1970), Knox-Thompson (Knox and Thompson 1974), and Weigelt's speckle masking (Lohmann et al. 1983). However, ordinary shift-and-add can only reconstruct images which have isolated bright points, while Knox-Thompson and Weigelt's methods in principle allow arbitrary objects to be reconstructed. The generalised shift-and-add principle may eventually allow arbitrary objects to be reconstructed in a less noise-sensitive manner. Generalised shift-and-add significantly enlarges the class of objects which can be reconstructed by the ordinary shift-and-add principle.

Three special cases of the generalised shift-and-add principle are identified. These are called 'correlation', 'inverse', and 'differential' shift-and-add. To implement these algorithms an ensemble of speckle images is transformed by cross-correlation, inverse filtering, and differentiation respectively. Usually differential shift-and-add is performed in conjunction with either correlation or inverse shift-and-add. The corresponding transformation is then a differentiation followed by cross-correlation or inverse filtering.

Correlation shift-and-add can handle a wide class of images but tends to reconstruct them somewhat distortedly. For instance, for an object

consisting of several isolated components (possibly extended), the locations and width of the components can usually be obtained but the relative amplitudes of the components tend to be closer to unity than they actually should be. In situations where only the number of components in the object is required, this is an advantage since the components are clearly identified.

Inverse shift-and-add reconstructs most objects which are dominated by several bright isolated points. The reconstructed amplitude tends to be more faithful than that of correlation shift-and-add.

Differential shift-and-add in conjunction with inverse or differential shift-and-add allows a significant proportion of one-dimensional piecewise constant extended objects to be reconstructed. It is anticipated that similar results may be obtained in two dimensions.

In this thesis relatively simple objects have been studied thoroughly. Hopefully future research workers in this field will investigate more complicated objects of the kind which figure in practical applications.

More research could be carried out on the model of shift-and-add presented in ¶6.7. An iterative solution of the model may allow a more faithful estimate of the object to be obtained than is provided by the application of the extent constraint described in ¶7.4.1. ¶10.4 contains some detailed suggestions in this direction.

Building on the preliminary experience discussed in ¶10.2, inverse and differential shift-and-add should be investigated in two dimensions and tested on experimental data.

The general formulation presented in Chapter 9 of shift-and-add shows that the transform technique may also be applied to correlate-and-add (¶6.5)). In correlate-and-add the threshold determines the level of ghosting. As the threshold tends to zero, the reconstruction tends to the autocorrelation. However, if the transform technique was applied to correlate-and-add then the threshold could approach zero without ghosting occurring. Since a low threshold corresponds to a low contamination in the correlate-and-add image, the transform technique is just as significant for correlate-and-add as it is for shift-and-add.

In the author's opinion, the most useful extension of the research reported here would be the application of generalised shift-and-add to more complicated objects. The author believes that generalised shift-and-add

imaging promises to play a significant role in CAI systems operating in a distorting medium.

A. INTERCONNECTIONS FOR THE ULTRASONIC SCANNER

The electronic instrumentation for the scanner described in ¶5.2 resides in a rack and on a gantry mounted on top of the tank (Fig. A.1).

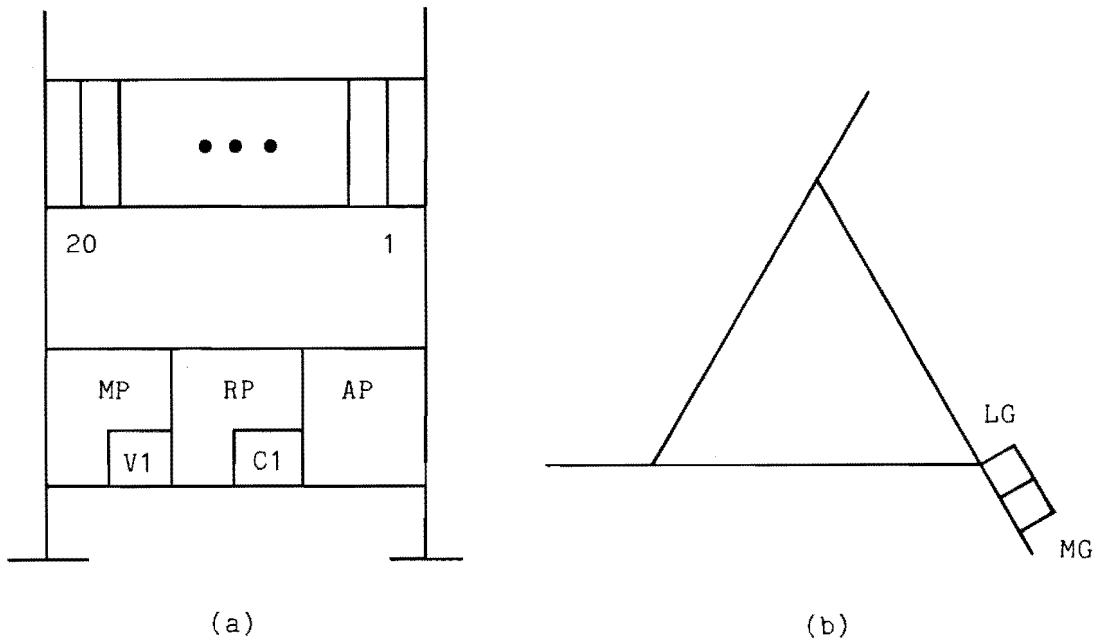


Fig. A.1 Layout of scanner instrumentation.
(a) Rear view of electronics rack.
(b) Top view of gantry.
The labelled components are identified in the text.

The rack has one row with slots for twenty circuit boards and a second row containing three power supplies. The slots are numbered as shown in Fig. A.1(a) and the supplies are labelled MP, RP, AP standing for main, radius, and angle power supplies respectively. A capacitor C1 and a voltage regulator V1 are mounted at the back of the second row along with some additional components which are shown in Fig. A.2. The electronics in the rack and on the gantry are connected by two cables which plug into the sockets mounted on the gantry. These sockets are marked LG and MG in Fig. A.1(b). Signals associated with the limit switches and drive waveforms for the stepper motors are carried by the LG and MG sockets respectively. Not

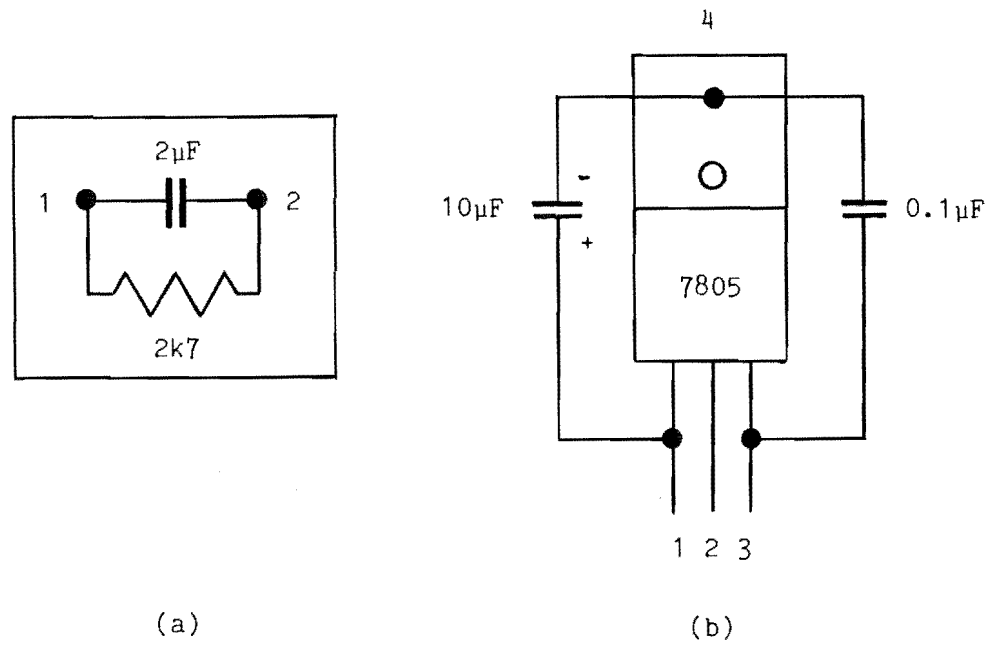


Fig. A.2 Node numbers for external circuitry as viewed from rear of electronics rack.
(a) The capacitor C1.
(b) The voltage regulator V1. Node 4 is the case of the regulator. Nodes 1, 2 and 3 are the input, ground and output respectively.

all of the twenty slots are occupied by circuit boards. The slots which are in use and the functions of the circuit boards in them are shown in Table A.1.

The wiring connecting the components shown in Fig. A.1 may be documented by assigning a symbol to every node and then for each node listing the nodes connected to it. The node identification for C1 and V1 is shown in Fig. A.2 and that for the gantry plugs LG and MG in Fig. A.3. The nodes for slots 1 to 20 and the power supplies are numbered on the scanner itself. The symbol for a node is formed by following the component symbol with a hyphen and the node number, e.g. the case of the voltage regulator is written V1-4 and the 2nd pin of the 3rd slot is 3-2. In some cases the connection between nodes takes the form of an electrical component other than a wire. The component is shown in the list, by enclosing it in

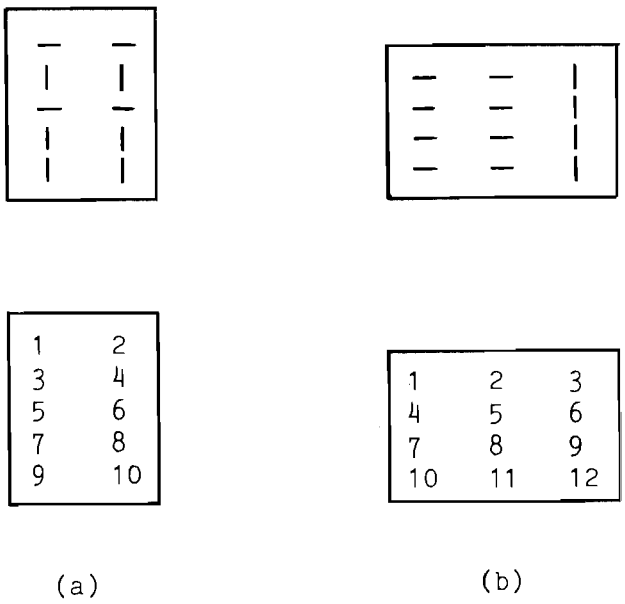


Fig. A.3 Gantry socket node identification.
(a) Limit switch socket LG.
(b) Motor plug MG.
The top row shows the physical appearance of the gantry sockets from above and the bottom row shows the node number assigned to each connection.

parentheses. For instance, suppose that 2-3 is connected to 4-5 by a 67Ω resistor. This would be written as (67Ω)4-5 in the list for 2-3. A list of node connections for the scanner is provided in Table A.2. Since each connection is between two nodes, it appears in two lists. This property of the node lists enables Table A.2 to be checked for consistency and this has been done.

In addition to the node lists in Table A.2, the signal present at each node is identified. Some of these signals are described in Table 5.1 and the remainder are described in Table A.3. The coaxial cables listed in Table A.4 are connected to the nodes to allow the signals to be monitored on an oscilloscope or connected to other instruments. For signals consisting of several bits, each bit is numbered as in Fig. A.4. Bit 7 of the ANGLE signal is written ANGLE 7, for instance. A cable, referred to as

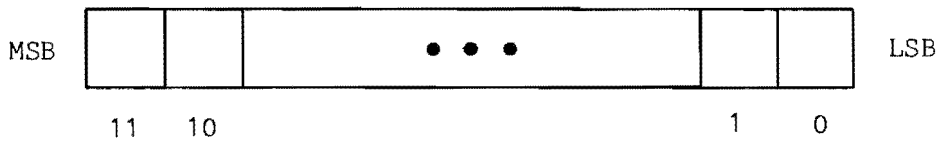


Fig. A.4 Bit numbering. The least significant bit (LSB) is bit 0. The most significant bit (MSB) of a 12-bit word is bit 11.

the 'A/D' cable, is attached to the board in slot 14 and carries signals between the scanner and the computer. The signals present on this cable are listed in Table A.5.

Table A.1 Scanner circuit boards.

Slot Number	Circuit Board Function
1	Position control system for radial motion (¶5.2.4).
2	Position control system for angular motion (¶5.2.4).
3	A cable is attached to the board which carries signals associated with the position control between the scanner and the computer.
4	Interrupt logic.
6	Timing circuitry.
8	5 V Power supply.
13	Reference path gain control.
16	A/D converter. A cable is attached to the board which carries signals including the output from the A/D converter between the scanner and the computer.
17	±12 V Power supply.
20	Scattered path gain control.

Slot 1 (Radius Position Control)		
Pin	To	Signal
1	2-1,1-32	Digital Ground
2	3-19	RADIUS 8/
3	3-20	RADIUS 9/
4	3-21	RADIUS 10/
5	3-22	RADIUS 11/
6	3-23	RADIUS 4/
7	3-24	RADIUS 5/
8	3-25	RADIUS 6/
9	3-26	RADIUS 7/
10	3-27	RADIUS 0/
11	3-28	RADIUS 1/
12	3-29	RADIUS 2/
13	3-30	RADIUS 3/
14		
15	2-15	
16	LG-6,4-30	SET_PHOTO_RADIUS/
17		
18	2-18,LG-4,LG-8	Ground
19	4-31,LG-5	CLR_PHOTO_RADIUS/
20	4-19	PHASE_ERR_RADIUS
21	4-22	
22	4-25	OVERFLOW_RADIUS/
23	4-27	UNDERFLOW_RADIUS/
24		
25		
26	RP-8	PHI1_RADIUS
27		
28		
29	RP-16	PHI2_RADIUS
30	RP-15	
31	2-31,3-15	5 V
32	1-1,2-32,3-31	Digital Ground

Slot 2 (Angle Position Control)		
Pin	To	Signal
1	1-1,3-32	Digital Ground
2	3-1	ANGLE 8/
3	3-2	ANGLE 9/
4	3-3	ANGLE 10/
5	3-4	ANGLE 11/
6	3-5	ANGLE 4/
7	3-6	ANGLE 5/
8	3-7	ANGLE 6/
9	3-8	ANGLE 7/
10	3-9	ANGLE 0/
11	3-10	ANGLE 1/
12	3-11	ANGLE 2/
13	3-12	ANGLE 3/
14		
15	1-15,4-16	
16	4-28,LG-2	SET_PHOTO_ANGLE/
17		
18	1-18,4-32,LG-4,LG-8	Ground
19	4-29,LG-1	CLR_PHOTO_ANGLE/
20	4-18	PHASE_ERR_ANGLE
21	4-20	
22	4-21	OVERFLOW_ANGLE/
23	4-23	UNDERFLOW_ANGLE/
24		
25		
26	AP-8	PHI1_ANGLE
27		
28		
29	AP-16	PHI2_ANGLE
30	AP-15	
31	1-31,8-3	5 V
32	8-2,1-32	Digital Ground

Table A.2 (cont.)

Slot 3 (Position Control Cable)		
Pin	To	Signal
1	2-2	ANGLE 8/
2	2-3	ANGLE 9/
3	2-4	ANGLE 10/
4	2-5	ANGLE 11/
5	2-6	ANGLE 4/
6	2-7	ANGLE 5/
7	2-8	ANGLE 6/
8	2-9	ANGLE 7/
9	2-10	ANGLE 0/
10	2-11	ANGLE 1/
11	2-12	ANGLE 2/
12	2-13	ANGLE 3/
13		
14		
15	4-2,1-31	5 V
16		
17		
18	4-17	SETUP/
19	1-2	RADIUS 8/
20	1-3	RADIUS 9/
21	1-4	RADIUS 10/
22	1-5	RADIUS 11/
23	1-6	RADIUS 4/
24	1-7	RADIUS 5/
25	1-8	RADIUS 6/
26	1-9	RADIUS 7/
27	1-10	RADIUS 0/
28	1-11	RADIUS 1/
29	1-12	RADIUS 2/
30	1-13	RADIUS 3/
31	3-32,1-32	Digital Ground
32	3-31	Digital Ground

Slot 4 (Interrupt Logic)		
Pin	To	Signal
1	4-2	5 V
2	4-1,3-15,	5 V
3	6-19,3-32	
4	6-28	MCLOCK
5	16-5	ADCLOCK
6	16-8	SWEEP_RQST/
7	16-7	INTERRUPT/
8	16-10	SWEEP/
9		
10	16-6	MOTOR/
11	AP-15	
12	AP-7	
13	6-1	
14	6-4	
15	6-5,20-7	
16	2-15	
17	3-18	SETUP/
18	2-20	PHASE_ERR_ANGLE
19	1-20	PHASE_ERR_RADIUS
20	2-21	
21	2-22	OVERFLOW_ANGLE/
22	1-21	
23	2-23	UNDERFLOW_ANGLE/
24		
25	1-22	OVERFLOW_RADIUS/
26	6-6	FAULT
27	1-23	UNDERFLOW_RADIUS/
28	2-16	SET_PHOTO_ANGLE/
29	2-19	CLR_PHOTO_ANGLE/
30	1-16	SET_PHOTO_RADIUS/
31	1-19	CLR_PHOTO_RADIUS/
32	2-18	Digital Ground

Table A.2 (cont.)

Slot 6 (Timing Circuitry)		
Pin	To	Signal
1	4-13	
2	20-6	
3	4-14	
4	4-15	
5	4-26,16-4	FAULT
6		
7		
8		
9		
10		
11	20-2	
12		
13		
14		
15		
16	8-4	
17		
18	8-1,4-3	
19		
20		
21		
22		
23		
24		
25		
26	4-4	MCLOCK
27		
28		
29		
30		
31		
32		

Slot 8 (5 V Power Supply)		
Pin	To	Signal
1	8-2,6-19,MP-13	
2	8-1,2-32	
3	LG-3,2-31	5 V
4	6-17	
5		
6		
7	8-18	
8	8-7,(fuse)MP-4	

Table A.2 (cont.)

Slot 13 (Reference Path Gain Control)		
Pin	To	Signal
1		
2		
3		
4		
5		
6		
7		
8	C-braid	
9		REF_MULT_OUT_LP
10	20-15	SCAT_MULT_OUT
11	(56Ω)13-7	SUM_MULT_OUT_LP
12		
13		
14		
15		
16		
17	(56Ω)13-11,C-core	SUM_MULT_OUT
18		
19		
20	20-20	TX_REF
21		
22		
23	17-5	+12 V
24		
25	17-1	12 V Ground
26		
27	17-4	-12 V
28		
29		REF_RF_SLEW_LIMITED
30		
31		REF_RF_BUFFERED
32		

Slot 16 (A/D Converter)		
Pin	To	Signal
1	MP-2	
2		
3	MP-10	
4	6-6	FAULT
5	4-5	ADCLOCK
6	4-10	MOTOR/
7	4-7	INTERRUPT/
8	4-6	SWEEP_RQST/
9		
10	4-8	SWEEP/
11		
12		
13		+15 V
14		
15		-15 V
16		
17		
18		
19		A/D Input
20		
21		20 V Ground
22		
23		
24		
25	17-1	20 V Ground
26		
27	17-2	+20 V
28		
29	17-3	-20 V
30		
31		
32		

Table A.2 (cont.)

Slot 17 (±12 V Power Supply)		
Pin	To	Signal
1	16-25,13-25,MP-3,20-25	12 V Ground
2	16-27,MP-11	
3	16-29,MP-9	
4	13-27,20-27	-12 V
5	13-23,20-23	+12 V
6		

Slot 20 (Scattered Path Gain Control)		
Pin	To	Signal
1		
2	6-12	
3		
4		
5		
6	6-2	
7	4-15	
8	D-braid	
9	(56Ω)20-15	SCAT_MULT_LP_OUT
10		
11		
12		
13		
14		
15	13-10,(56Ω)20-9,D-core	
16		
17		
18		
19	T-braid	
20	13-20,T-core	TX_REF
21		
22		
23	17-5	+12 V
24		
25	17-1	12 V Ground
26		
27	17-4	-12 V
28		
29	E-core	SCAT_RF_SLEW_LIMITED
30	E-braid	
31	F-core	SCAT_RF_BUFFERED
32	F-braid	

Table A.2 (cont.)

Slot AP (Angle Power Supply)		
Pin	To	Signal
1	RP-1	
2		
3	MG-7	
4	MG-11	
5	AP-13,C1-1,AP-13	
6	AP-14,C1-2	
7	4-12,RP-7	
8	2-26	PHI1_ANGLE
9		
10		
11	MG-8	
12	MG-10,AP-13,AP-5	
13	AP-12,AP-5	
14	AP-6	
15	2-30,4-11,RP-15	
16	2-29	PHI2_ANGLE

Slot MP (Main Power Supply)		
Pin	To	Signal
1	MP-3	
2	16-1	
3	MP-1,17-1	
4		
5	MP-13	
6	MP-14	
7	MP-15	
8	MP-16	
9	17-3	
10	16-3	
11	17-2	
12		
13	V1-4,8-1,MP-5	
14	MP-6,(diode)V1-1,(fuse)8-8	
15	MP-chassis,RP-6,MP-7	
16	MP-8,(fuse)C1-1	
chassis	MP-5	

Slot RP (Radius Power Supply)		
Pin	To	Signal
1	AP-1,V1-3	
2		
3	MG-1	
4	MG-5	
5	RP-13	
6	RP-14,MP-15	
7	AP-7	
8	1-26	PHI1_RADIUS
9		
10		
11	MG-2	
12	MG-4	
13	RP-5,C1-1	
14	RP-6,C1-2	
15	AP-15,1-30	
16	1-29	PHI2_RADIUS

Table A.2 (cont.)

Socket LG (Limit Switch Gantry Plug)		
Pin	To	Signal
1	2-19	CLR_PHOTO_ANGLE/
2	2-16	SET_PHOTO_ANGLE/
3	8-3,(0.1μF)LG-4	5 V
4	2-18,(0.1μF)LG-3	Digital Ground
5	1-19	CLR_PHOTO_RADIUS/
6	1-16	SET_PHOTO_RADIUS/
7		
8	1-18	Ground
9		
10		

Socket MG (Motors Gantry Plug)		
Pin	To	Signal
1	RP-3	
2	RP-11	
3		
4	RP-12	
5	RP-4	
6		
7	AP-3	
8	AP-11	
9		
10	AP-12	
11	AP-4	
12		

Table A.3 Signal Descriptions.

Signal	Description
SET_PHOTO_ANGLE	True when limit switch at upper end of angle travel is reached.
CLR_PHOTO_ANGLE	True when limit switch at lower end of angle travel is reached.
OVERFLOW_ANGLE	Indicates that the angle current position counter has overflowed.
UNDERFLOW_ANGLE	Indicates that the angle current position counter has underflowed.
PHI1_ANGLE PHI2_ANGLE	A biphase pair used to drive the angle stepper motor forward or backward depending on the phase relationship of the signals.
PHASE_ERR_ANGLE	Indicates that the biphase pair is incorrect in some way.
SET_PHOTO_RADIUS CLR_PHOTO_RADIUS OVERFLOW_RADIUS UNDERFLOW_RADIUS PHI1_RADIUS PHI2_RADIUS PHASE_ERR_RADIUS	These signals are for the radius motion and correspond to the above signals for the angle motion.
MCLOCK	Master clock signal with the frequency given by (5.9).
ADCLOCK	Samples are taken on the positive going edges of this signal and then are digitised.
SWEEP	This signal goes high at t = 0, triggering the sweep. The sweep is reset when the signal goes low.

Table A.3 (cont.)

Signal	Description
TX_REF	Reference signal proportional to the transmitter output ($a_T(t)$ as defined in ¶5.2.5).
SCAT_MULT_OUT_LP	Low pass filtered output of multiplier for scattered path ($a_{TS}(t)$).
SCAT_RF_SLEW_LIMIT	Slew limited version of the RF signal from transducer S.
SCAT_RF_BUFFERED	Buffered version of the RF signal from transducer S ($a_S(t)$).
REF_MULT_OUT_LP REF_RF_SLEW_LIMIT REF_RF_BUFFERED	These signals are for the reference path and correspond to those above for the scattered path.
SUM_MULT_OUT_LP	Sum of low pass filtered multiplier outputs ($a_{TSRL}(t)$).

Table A.4 Coaxial cables.

Cable Name	Signal
C	SUM_MULT_OUT
D	SCAT_MULT_OUT_LP
T	TX_REF
E	SCAT_RF_SLEW_LIMITED
F	SCAT_RF_BUFFERED

Table A.5 The signals present on the A/D cable.

Pin	Signal
1	Ground
2	A/D 11
3	DATA_READY
4	A/D 9
5	INTERRUPT
6	A/D 10
7	SWEEP_RQST/
8	A/D 8
9	FAULT
10	A/D 4
11	MOTORS
12	A/D 5
13	5 V
14	A/D 6
15	Spare
16	A/D 7
17	Ground
18	A/D 0
19	Spare
20	A/D 1
21	
22	A/D 2
23	
24	A/D 3
25	
26	Spare
27	
28	Spare
29	
30	Spare
31	Ground
32	Spare

B. THE CABLE BOX FOR THE SCANNER-VAX INTERFACE

The cables connecting the scanner to the VAX are terminated at sockets mounted in a box. The layout of the sockets is shown in Fig. B.1.

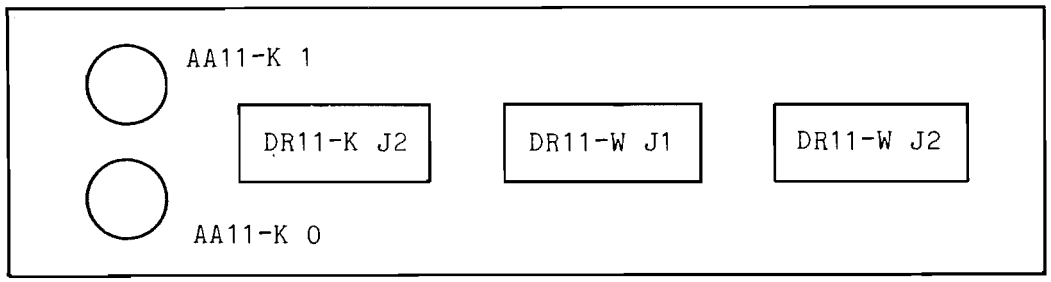


Fig. B.1 Elevation of box showing the socket layout.

Two 5-pin DIN sockets are used for the D/A cables associated with the AA11-K and three Berg connectors are used for the digital signals. The pin numbering for these sockets is shown in Fig. B.2. Table B.1 lists the signals assigned to each pin.

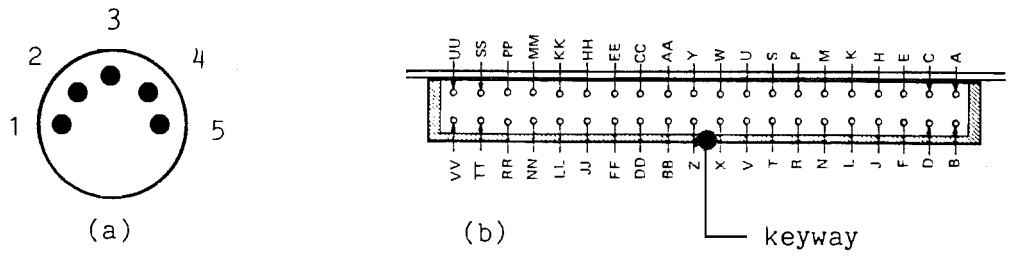


Fig. B.2 Pin numbering.
(a) 5-pin DIN socket. View shown is that seen when looking into socket from plug side.
(b) 40-way BERG connector. The connector is shown as it is seen when the box is viewed from above.

Table B.1 Signal assignments.

(a) AA11-K socket 0

Pin	Signal
1	Scattered path gain control voltage
2	Ground
3	Shield

(b) AA11-K socket 1

Pin	Signal
1	Reference path gain control voltage
2	Ground
3	Shield

(c) DR11-K J2 connector

DR11-K Signal	Pin	Scanner Signal
INIT	A	
EXT DATA ACC	C	5V
INT HI DATA RDY	E	
INT LO DATA RDY	H	
OUT00	K	RADIUS 0/
OUT01	M	RADIUS 1/
OUT02	P	RADIUS 2/
OUT03	S	RADIUS 3/
OUT04	U	RADIUS 4/
OUT05	W	RADIUS 5/
OUT06	Y	RADIUS 6/
OUT07	AA	RADIUS 7/
OUT08	CC	RADIUS 8/
OUT09	EE	RADIUS 9/
OUT10	HH	RADIUS 10/
OUT11	KK	RADIUS 11/
OUT12	MM	
OUT13	PP	
OUT14	SS	
OUT15	UU	

Scanner Signal	Pin	DR11-K Signal
	B	GND
	D	GND
	F	GND
	J	GND
	L	GND
	N	GND
	R	GND
	T	GND
	V	GND
	X	GND
	Z	GND
	B	GND
	DD	GND
	FF	GND
	JJ	GND
	LL	GND
	NN	GND
	RR	GND
	TT	GND
	VV	GND

(d) DR11-W J1 connector

DR11-W Signal	Pin	Scanner Signal
GND	A	
GND	C	
GND	E	
GND	H	
D5 BURST RQ L	K	5V
GND	M	
GND	P	
GND	S	
GND	U	
GND	W	
GND	Y	
GND	AA	
D10 DO 07 H	CC	ANGLE 7/
D10 DO 06 H	EE	ANGLE 6/
D10 DO 05 H	HH	ANGLE 5/
D10 DO 04 H	KK	ANGLE 4/
D10 DO 03 H	MM	ANGLE 3/
D10 DO 02 H	PP	ANGLE 2/
D10 DO 01 H	SS	ANGLE 1/
D10 DO 00 H	UU	ANGLE 0/

Scanner Signal	Pin	DR11-W Signal
DATA_READY	B	D5 CYCLE RQ A H
	D	D5 ACLO FNCT 2 H
	F	D5 READY H
5V	J	D5 WC INC ENB H
MOTOR	L	D5 STATUS A H
	N	D5 INIT H
FAULT	R	D5 STATUS B H
	T	D5 STATUS C H
	V	D5 STATUS C H
	X	D5 END CYCLE H
DATA_READY	Z	D5 CYCLE RQ B H
	BB	GND
ANGLE 8/	DD	D10 DO 08 H
ANGLE 9/	FF	D10 DO 09 H
ANGLE 10/	JJ	D10 DO 10 H
ANGLE 11/	LL	D10 DO 11 H
	NN	D10 DO 12 H
	RR	D10 DO 13 H
	TT	D10 DO 14 H
	VV	D10 DO 15 H

(e) DR11-W J2 connector

DR11-W Signal	Pin	Scanner Signal
GND	A	
GND	C	
GND	E	
GND	H	
D5 FNCT 3 H	K	
GND	M	
GND	P	
GND	S	
GND	U	
GND	W	
GND	Y	
GND	AA	
D4 DI 07 H	CC	A/D 7
D4 DI 06 H	EE	A/D 6
D4 DI 05 H	HH	A/D 5
D4 DI 04 H	KK	A/D 4
D4 DI 03 H	MM	A/D 3
D4 DI 02 H	PP	A/D 2
D4 DI 01 H	SS	A/D 1
D4 DI 00 H	UU	A/D 0

Scanner Signal	Pin	DR11-W Signal
	B	D5 BUSY H
INTERRUPT	D	D5 ATTN H
0V	F	D5 A00 H
5V	J	D5 BA INC ENB H
	L	D5 FNCT 3 H
0V	N	D5 C0 CNTL H
SWEEP_RQST/	R	D5 FNCT 2 H
5V	T	D5 C1 CNTL H
SETUP/	V	D5 FNCT 1 H
	X	D5 GO H
	Z	GND
	BB	GND
A/D 8	DD	D4 DI 08 H
A/D 9	FF	D4 DI 09 H
A/D 10	JJ	D4 DI 10 H
A/D 11	LL	D4 DI 11 H
	NN	D4 DI 12 H
	RR	D4 DI 13 H
	TT	D4 DI 14 H
	VV	D4 DI 15 H

C. HELP DISPLAYS FOR SOME GEM COMMAND MODULES

Figs. C.1, C.2, C.3, C.4, C.5, and C.6 show the responses to the Help command for the image>, mop>, mio>, matrix>, and field> GEM command modules (¶4.8).

```
Image commands:

        Centroid                image   point
        Convolve                A       B/Extend=ext_x ext_y    C
        Correlate               A       B/Extend=ext_x ext_y    C
        Differentiate/X(I)/Y    from/Extend=ext_x ext_y to (inplace ok)
        Format/format_mode      image   put
        Fourier/Forward         image   matrix
        Fourier/Inverse         matrix  image
                /[No]Swap/[No]Window
                the default is /forward/swap/nowindow
        Histogram
                one/Amplitude=one_range                         histogram
        Integrate/X(I)/Y        from/Extend=ext_x ext_y to (inplace ok)
        Joint/Displacement=field
                one/Amplitude=one_range two/Amplitude=two_range histogram
        histogram =     x_hist_joint(I A1)/[No]Auto(I)
                        y_hist_joint(I A2)/[No]Auto(I)
                        hist_image
                                /[No]Zero(I)/Normalise=hist_norm(I XY)
                                /Area=area
        Moment                  image   to_scalar
                /Centroid[=point](D)/Origin=point/Power=scalar(I 1)
        Plot    x_par/x_options y_par/y_options         dest
                /X              x_par   = x_val
                /FX             x_par   = x_val x_fn
                /X_sweep(D)     x_par   = x_range
                - similarly for /Y/FY/Y_sweep
        Sample  /Image=image
                /From_Field=from_field/From_Matrix=matrix
                /Size=from_size/Origin=from_origin
                /Matrix(D)      to_size to_origin
                /Transfer       to_field to_size to_origin shift
                /Field          to_size to_origin to_field
                /Index          point index
                /Point          index point
                /Delta          delta
        XYA     x       y       a       image
        format_mode     =       /[No(I)]Get/[No]Symbol(I)/[No]Value(I)
                                /Precision=prec_vec(I 2 2)
        put_area        = {     Print(I)        just (left I)   print
                                Area            area
                                Field           field           device}
        put_form        = {     Print(I)        just (left I)   print
                                Area            area
                                Device_mask     device}
                                Symbol          symbol}
```

Fig. C.1 Help display for image> command module.

```
Matrix operations available are:

        Add             A       B       result
        Average         matrix  to_scalar
        Constant        scalar  matrix
        Copy            from    to
        Divide          A       B       result
        Flip1           matrix  (D)
        Flip2           from    to
        Fourier         matrix
                /Forward/Inverse/(No)Swap/(No)Window
                the default is /forward/swap/nowindow
        Function        from    fn      to
        Gaussian        matrix
                /Centre=x,y/Amplitude=r4/Base=r4/Radius=r4
        Insert          from    {Previous Zero} {Real Imaginary} to
        Linear          plane_vector    matrix
        Matrix          from    arith   scalar  to
        Minmax/Noprint  matrix  min_vec min_sca max_vec max_sca
        Minmax/Print(D) matrix  just/print
                /No(D)Absolute, omit vectors for complex matrices
        Multiply        A       B       result
        Normalise/Maximum=scalar        from    to
        Phase           from    ref_vec sec_vec to
        Quantise        floating        quantised
        Float           quantised       floating
                /dtype/Help/Range=range/Uniform(D)[=range]
        Optimise/Matrix=fac/vector=fac
        Reverse         from    to
                /XY(D)/X/Y/None (inplace o.k.)
        Sample[/Fourier(D)/Delete]      from    to
                n.b. Samp/fou leaves swapped FT in from
        Scalar          scalar  arith   from    to
        Subtract        A       B       result
        Sum/Zero(D)/Nozero      matrix  scalar
        Swap            from    to
        Threshold       from    centre  width   to
                /Discontinuous(I)/Continuous/Power=power (I 1)
        Transpose       from    to (inplace o.k.)
        Window          from wind_x wx [,wind_y wy] to
                ,wind_y wy if omitted is assumed to be as for x
                wind_z is one of {Hanning Triangle}
        Wrap            from    shift_vector    to
```

Fig. C.2 Help display for mop> command module.

```
Matrix input/output commands are:

        Draw    /Device=graphics/Spiral(I)/Scplot/Star
                /Radius=r4_sca(I 0.8)/[No]Fraction(I)
                /Scale[=(X,Y)]  matrix  field
        Format/format_mode      matrix                  put
        Shade   /Device=graphics/Colour/Plot(I)
                /Scale[=(X,Y)]  matrix  field
        n.b.    omit matrix for /Scale
        Shade[/unit]            matrix  title
        Shade[/Device=(Fads,Vips)]      matrix  pos
                1) invert ignored for Fads and graphics
                2) window not required for byte arrays
        Test                    matrix

        Structure               struct  attribute thing
        Tabulate                struct  cols title just area

        Values[/unit]           matrix title precision
        Values[/File=filename]  matrix title precision width

        (Dump,Load)/Title=text[/quantiser]      matrix /move
                /All(D)/Nosize(D)/Noorigin(D)/Size=(sx sy)/Origin=(x1 y1)
        Plot    x_par/x_options y_par/y_options         dest
                /X              x_par   = x_val
                /FX             x_par   = x_val x_fn
                /X_sweep(D)     x_par   = x_range
                - similarly for /Y/FY/Y_sweep

        Contour                 matrix  vector  dest
                /Cubic=xinterp(I 1),yinterp(I 1)
                /Join=num(I 1000)/Label(I)/NoLabel
        Get                     matrix pos (gets sub-matrix from Fads)
        Type                    matrix width    /print

        graphics = Screen,Plotter,Gigi,invalid(I)
        unit     = Printer (width=132),Out(D),Log
        spiral>, window>
```

Fig. C.3 Help display for mio> command module.

```
Matrix commands are:

        Declare                 matrix mode[*length] sx sy nx ny x1 y1
                                default nx=sx, ny=sy, x1=1, y1=1
        Format/format_mode      matrix  put
        Release                 matrix
```

Fig. C.4 Help display for matrix> command module.

```
Field commands are:

        Copy                    from    to
        [No]Debug
        Declare                 field   field_lit
        {Dump,Load}/Title=text  field/move
        Format/format_mode      field   put
        Origin                  field
        Plot                    field   device
        Release                 field
        Restore
        Set                     field   field_lit
        Show                    field   /print
        Use                     field
        field_lit       = {= xor yor xs ys,? device}
```

Fig. C.5 Help display for field> command module.

REFERENCES

Many of the referenced papers are from the series Acoustical Holography/Imaging. Details of the twelve volumes of this series are:

Acoustical Holography 1, A.F. Metherell, H.M.A. El-Sum and L. Larmore Ed., Plenum Press, New York, 1969.

Acoustical Holography 2, A.F. Metherell and L. Larmore Ed., Plenum Press, New York, 1970.

Acoustical Holography 3, A.F. Metherell Ed., Plenum Press, New York, 1971.

Acoustical Holography 4, G. Wade Ed., Plenum Press, New York, 1972.

Acoustical Holography 5, P.S. Green Ed., Plenum Press, New York, 1974.

Acoustical Holography 6, N. Booth Ed., Plenum Press, New York, 1975.

Acoustical Holography 7, L.W. Kessler Ed., Plenum Press, New York, 1977.

Acoustical Imaging 8, A.F. Metherell Ed., Plenum Press, New York, 1980.

Acoustical Imaging 9, K.Y. Wang Ed., Plenum Press, New York, 1980.

Acoustical Imaging 10, P. Alais and A.F. Metherell Ed., Plenum Press, New York, 1982.

Acoustical Imaging 11, J.P. Powers Ed., Plenum Press, New York, 1982.

Acoustical Imaging 12, E.A. Ash and C.R. Hill Eds., Plenum Press, New York, 1982.

Abbott J.G. and Thurstone F.L. 1979. "Acoustic speckle: Theory and experimental analysis", Ultrasonic Imaging 1, 303-324.

Abramowitz M. and Stegun I.A. 1965. "Handbook of Mathematical Functions", Dover, New York.

Adams M.F. and Anderson A.P. 1982. "Tomography from multiview diffraction data: Comparision with image reconstruction from projections", Acoustical Imaging 10, 365-380.

Alais P. and Richard B. 1980. "Progress in Fresnel imaging", Acoustical Imaging 8, 137-148.

Andersen A.H. and Kak A.C. 1984. "Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm", Ultrasonic Imaging 6, 81-94.

Arditi M., Foster F.S. and Hunt J.W. 1981. "Transient fields of concave annular arrays", Ultrasonic Imaging 3, 37-63.

Assenza D. and Pappalardo M. 1980. "Echographic imaging with dynamically focused insonification", Ultrasonics 18, 38-42.

Auld B.A. 1973. "Acoustic Fields and Waves in Solids", Vol. 1, J. Wiley and Sons, New York.

Azimi M. and Kak A. 1983. "Distortion in diffraction tomography caused by multiple scattering", IEEE Trans. MI-2, 176-195.

Bagnuolo W.G. (Jr.) 1982. "The application of Bates' algorithm to binary stars", Mon. Not. R. Astron. Soc. 200, 1113-1122.

Bates R.H.T. 1967. "The point matching method for interior and exterior two-dimensional boundary value problems", IEEE Trans. MTT-15, 185-187.

Bates R.H.T. 1969. "Contributions to the theory of the intensity interferometer", Mon. Not. R. Astron. Soc. 142, 413-428.

Bates R.H.T. and Ng F.L. 1972. "Polarisation-source formulation of electromagnetism and dielectric-loaded waveguides", Proc. IEEE 119, 1568-1574.

Bates R.H.T., James J.R., Gallett I.N.L and Millar R.F. 1973. "An overview of point matching", Radio Electron. Eng. 43, 193-200.

Bates R.H.T. 1974. "New justification for physical optics and the aperture-field method", in Proc. 20th AGAARD Meeting on Electromagnetic Wave Propagation Involving Irregular Surfaces and Inhomogeneous Media (The Hague, The Netherlands), Mar. 25-29, 36-1 - 36-7, AGAARD Conf. Publication No. CPP-144.

Bates R.H.T. 1975. "Analytic constraints on electromagnetic field computations", IEEE Trans. MTT-23, 605-623.

Bates R.H.T., Boerner W.M. and Dunlop G.R. 1976. "An extended Rytov approximation and its significance for remote sensing and inverse scattering", Opt. Commun. 18, 421-423.

Bates R.H.T. and Dunlop G.R. 1977. "Inverse scattering and tomography", Ultrasonics International 1977 (Conference Proceedings), I.P.C. Science and Technology Press, Guildford U.K., 104-110.

Bates R.H.T and Wall D.J.N. 1977a. "Null field approach to scalar diffraction: I General method", Phil. Trans. R. Soc. Lond. A. 287, 45-78.

Bates R.H.T and Wall D.J.N. 1977b. "Null field approach to scalar diffraction: II Approximate methods", Phil. Trans. R. Soc. Lond. A. 287, 79-95.

Bates R.H.T and Wall D.J.N. 1977c. "Null field approach to scalar diffraction: III Inverse methods", Phil. Trans. R. Soc. Lond. A. 287, 97-114.

Bates R.H.T. and Cady F.M. 1980. "Towards true imaging by wideband speckle interferometry", Opt. Commun. 32, 365-369.

Bates R.H.T. and McKinnon G.C. 1980. "Possible applications of inverse scattering concepts to ultrasonic imaging", Research Techniques in Non-Destructive Testing 4, R.S. Sharp Ed., Academic Press, London, Ch. 3, 71-85.

Bates R.H.T. 1981. "Phase recovery, speckle and image reconstruction", Proc. SPIE Conf. Transformations in Optical Signal Processing 373, 191-196.

Bates R.H.T. and Millane R.P. 1981. "Time domain approach to inverse scattering", IEEE Trans AP-29, 359-363.

Bates R.H.T. and Robinson B.S. 1981. "Ultrasonic transmission speckle imaging", Ultrasonic Imaging 3, 378-394.

Bates R.H.T. 1982. "Astronomical speckle imaging", Phys. Rep. 90, 203-297.

Bates R.H.T and Fright W.R. 1982. "Towards imaging with a speckle-interferometric optical synthesis telescope", Mon. Not. R. Astron. Soc. 198, 1017-1031.

Bates R.H.T., Garden K.L. and Peters T.M. 1983. "An overview of computerized tomography with emphasis on future developments", Proc. IEEE 71, 356-372.

Bates R.H.T. and Minard R.A. 1983. "Some new approaches to inverse scattering", Proc. SPIE Conf. Inverse Optics, 413, 56-60.

Bates R.H.T. 1984. "Full-wave computed tomography. Part 1: Fundamental theory", IEE Proc. A 131, 610-615.

Bates R.H.T. and Fright W.R. 1984. "Reconstructing images from their Fourier intensities" in Advances in Computer Vision and Image Processing, T.S. Huang Ed., JAI Press, Volume I, in press.

Bates R.H.T. and McDonnell M.J. 1985. "Image restoration and Reconstruction", Oxford University Press.

Beckmann P. and Spizzichino A. 1963. "The Scattering of Electromagnetic Waves from Rough Surfaces", Pergamon, Oxford.

Blake L.V. 1966. "Antennas", J. Wiley and Sons, New York.

Bom N., Lancee C.T., van Zwieten G., Kloster F.E. and Roelandt J. 1973. "Multiscan echocardiography: (1) Technical description", Circulation 48, 1066-1074.

Born M. and Wolf E. 1965. "Principles of Optics", Pergamon Press, Oxford, 3rd Edition.

Bowman J.J., Senior T.B.A. and Uslenghi P.L.E. (Eds.) 1969. "Electromagnetic and Acoustic Scattering by Simple Shapes", North-Holland, Amsterdam.

Bracewell R.N. 1978. "The Fourier Transform and its Applications", McGraw-Hill, New York, 2nd Edition.

Bradley G.J. and Cracknell A.P. 1972. "The Mathematical Theory of the Symmetry of Solids", Clarendon Press, Oxford.

Brigham E.O. 1974. "The Fast Fourier Transform", Prentice-Hall, Englewood Cliffs, N.J.

Brisley W. 1973. "A Basis for Linear Algebra", J. Wiley and Sons, Sydney.

Bruneel C., Delannoy B., Torguet R., Bridoux E. and Lasota H. 1979. "Electrical coupling effects in an ultrasonic transducer array", Ultrasonics 17, 255-260.

Burckhardt C.B., Grandchamp P.A. and Hoffmann H. 1974. "Methods for increasing the lateral resolution of B-scan", Acoustical Holography 5, 391-414.

Burckhardt C.B. 1978. "Speckle in ultrasound B-mode scans", IEEE Trans. SU-25, 1-6.

Burkill J.C. and Burkill H. 1970. "A Second Course in Mathematical Analysis", Cambridge University Press, London.

Carson P.L., Oughton T.V., Hendee W.R. and Ahuja A.S. 1977. "Imaging soft tissue through bone with ultrasound transmission tomography by reconstruction", Med. Phys. 4, 302-309.

Cauvard P. and Hartemann P. 1982. "Ultrafast acoustical imaging with surface wave components", Acoustical Imaging 10, 17-22.

Chadan, K. and Sabattier P.C. 1977. "Inverse Problems in Quantum Scattering Theory", Springer-Verlag.

Chapelon J.Y., Cathignol D. and Fourcade C. 1979. "Improved ultrasonic sensitivity using pseudo-random binary-code phase-modulated signals", Ultrasonic Imaging 1, 255-264.

Chase R.C., Seguin F.H., Gerassimenko M. and Petrasso R. 1981. "Application of computerized tomography techniques to tokamak diagnostics", Optical Engineering 20, 486-492.

Chernov L.A. 1960. "Wave Propagation in a Random Media", McGraw-Hill, New York.

Clemmow P.C. 1966. "The Plane Wave Spectrum Representation of Electromagnetic Fields", Pergammon Press, Oxford.

Coes L. 1971. "Abrasives", Springer-Verlag, Wien.

Conte S.D. and de Boor C. 1972. "Elementary Numerical Analysis: An Algorithmic Approach", McGraw-Hill Kogakusha, Tokyo, 2nd Edition.

Cook C.E. and Bernfeld M. 1967. "Radar Signals", Academic Press, New York.

Cornbleet S. 1983. "Geometrical optics reviewed: A new light on an old subject", Proc. IEEE 71, 471-502.

Courant R. 1936. "Differential and integral calculus", J. Wiley and Sons, New York.

Cowley J.M. 1975. "Diffraction Physics", North-Holland, Amsterdam.

Crawford C.R. and Kak A.C. 1981. "Multipath Artifacts in Ultrasonic Transmission Tomography", School of Electrical Engineering, Purdue University, Tech. Rep. TR-EE-81-43.

Cromwell L., Weibell F.J. and Pfeiffer E.A. 1980. "Biomedical Instrumentation and Measurements", Prentice-Hall, Englewood Cliffs, N.J.

Daly P. 1974. "Polar geometry waveguides by finite-element methods", IEEE Trans. MTT-22, 202-209.

DeClercq A. and Maginness M.G. 1975. "Adaptive gain control for ultrasound imaging", IEEE Ultrasonics Symposium Proceedings, IEEE catalog no. 75 CHO 994-4SU, 59-63.

Devaney A.J. 1983. "A computer simulation study of diffraction tomography", IEEE Trans. BME-30, 337-386.

Devaney A.J. 1982. "A filtered backpropagation algorithm for diffraction tomography", Ultrasonic Imaging 4, 336-350.

Devaney A.J. and Beylkin G. 1984. "Diffraction tomography using arbitrary transmitter and receiver surfaces", Ultrasonic Imaging 6, 181-193.

Digital Equipment Corporation a. "DR11-K Interface User's Guide and Maintenance Manual", Order no. EK-DR11K-MM-001, Maynard, Massachusetts.

Digital Equipment Corporation b. "DR11-W Direct Memory Interface Module: User's Guide", Order no. EK-DR11W-UG-001, Maynard, Massachusetts.

Digital Equipment Corporation c. "LPA11-K Laboratory Peripheral Accelerator User's Guide", Order no. EK-LPA11-UG-001, Maynard, Massachusetts.

Digital Equipment Corporation d. "4 Channel D/A and Display Control User's Manual", Order no. EK-AA11K-TM-001, Maynard, Massachusetts.

Digital Equipment Corporation e. "Analog to Digital Converter User Manual", Order no. EK-AD11K-OP-002, Maynard, Massachusetts.

Digital Equipment Corporation 1978. "VAX-11 Software Handbook", Maynard, Massachusetts.

Digital Equipment Corporation 1979. "VAX-11 Architecture Handbook", Maynard, Massachusetts.

Digital Equipment Corporation 1980. "VAX Hardware Handbook", Maynard, Massachusetts.

Digital Equipment Corporation 1982a. "VAX-11 Symbolic Debugger Reference Manual", Order No. AA-D026D-TE, Maynard, Massachusetts.

Digital Equipment Corporation 1982b. "VAX-11 Utilities Reference Manual", Order No. AA-H781B-TE, Maynard, Massachusetts.

Digital Equipment Corporation 1982c. "VAX-11 Run-Time Library Reference Manual", Order No. AA-D036C-TE, Maynard, Massachusetts.

Digital Equipment Corporation 1982d. "VAX-11 Record Management Services Reference Manual", Order No. AA-D031D-TE, Maynard, Massachusetts.

Digital Equipment Corporation 1983. "VAX-11 I/O User's Guide", 2 volumes, Order Nos. AA-M540B-TE, AA-M541B-TE, Maynard, Massachusetts.

Dines K.A. and Kak A.C. 1979. "Ultrasonic attenuation tomography of soft tissues", Ultrasonic Imaging 1, 16-33.

Duck F.A., Johnson S.A., Greenleaf J.F. and Samayoa W.F. 1977. "Digital image focusing in the near field of a sampled acoustic aperture", Ultrasonics 15, 83-88.

Duerinckx A.J. 1981. "Modeling wavefronts from acoustic phased arrays by computer", IEEE Trans. BME-28, 221-234.

Dunlop G.R., Boerner W.M. and Bates R.H.T. 1976. "On an extended Rytov approximation and its comparison with the Born approximation", in Digest of Papers, AP-S International Symposium, IEEE catalog no. 76 CH 1121-3AP.

Dunlop G.R. 1978. "Ultrasonic Transmission Imaging", Ph.D. Thesis, University of Canterbury, New Zealand.

Dunn F. and Fry F.J. 1980. "Acoustic elliptic lenses: An historical note", J. Acoust. Soc Am. 68, 350.

Eaton M.D., Melen R.D. and Meindl J.D. 1980. "A flexible, real time system for experimentation in phased-array ultrasound imaging", Acoustical Imaging 8, 55-67.

Engan H., Ingebrigtsen K. and Ronnekleiv A. 1984. "On the design of high performance planar surfce acoustic wave convolvers", IEEE Trans. SU-31, 175-184.

Farrell E.J. 1981. "Tomographic imaging of attenuation with simulation correction for refraction", Ultrasonic Imaging 3, 144-163.

Finkelstein L. 1973. "Fundamental concepts of measurement" in Measurement and Instrumentation acta IMEKO 1973, Vol 1., Measurement Concepts, 11-28.

Foster F.S., Arditi M., Patterson M.S., Lee-Chahal D. and Hunt J.W. 1983. "Breast imaging with a conical transducer annular array hybrid scanner", Ultrasound in Med. & Biol. 9, 151-164.

Fraser J., Havlice J., Kino G., Leung W., Shaw H., Toda K., Waugh T., Winslow D. and Zitelli L. 1975. "An electronically focused two dimensional acoustic imaging system", Acoustical Holography 6, 275-304.

Gabor D. 1946. "Theory of communication", J. IEE (Lond.) 3, 429-441.

Gammell P.M. 1981. "Improved ultrasonic detection using the analytic signal magnitude", Ultrasonics 19, 73-76.

Garden K.L. 1984. "An overview of computed tomography", Ph.D. Thesis, University of Canterbury, New Zealand.

Glover G.H. and Sharp J.C. 1977. "Reconstruction of ultrasound propagation speed distributions in soft tissue: Time of flight tomography", IEEE Trans. SU-24, 229-234.

Goodman J.W. 1975. "Statistical properties of laser speckle", Laser Speckle and Related Phenomena, J.C. Dainty Ed., Springer-Verlag, Berlin, 9-77.

Gordon R., Bender R. and Herman G.T. 1970. "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography", J. Theor. Biol. 29, 471-481.

Gore J.C. and Leeman S. 1977. "Ultrasonic backscattering from human tissue: a realistic model", Phys. Med. Biol. 22, 317-326.

Green C.D. 1969. "Integral Equation Methods", Nelson, London.

Greenleaf J.F., Johnson S.A., Lee S.L., Herman G.T. and Wood E.H. 1974. "Algebraic reconstruction of spatial distributions of acoustic absorption within tissue from their two dimensional acoustic projections", Acoustical Holography 5, 591-603.

Greenleaf J.F., Johnson S.A., Samayoa W.F. and Duck F.A. 1975. "Algebraic reconstruction of spatial distributions of acoustical velocities in tissues from their time of flight profiles", Acoustical Holography 6, 71-90.

Greenleaf J.F., Johnson S.A., Bahn R.C., Rajagopalan B. and Kenue S. 1979. "Introduction to computed ultrasonic tomography", Computer Aided Tomography and Ultasonics in Medicine, J. Raviv, J.F. Greenleaf and G.T. Herman Ed., North-Holland, Amsterdam, 125-136.

Greenleaf J.F., Thomas P.J. and Rajagopalan B. 1982. "Effects of diffraction on ultrasonic computer-assisted tomography", Acoustical Imaging 11, 351-363.

Greenleaf J.F. and Chu A. 1984. "Multifrequency diffraction tomography", Acoustical Imaging 13 (submitted).

Greenleaf J.F., Robinson B.S., Kaveh M. and Thomas P. 1984. "Diffraction tomography of complex objects", 9th International Symposium on Ultrasonic Imaging and Tissue Characterisation 1984 (Abstract), in Ultrasonic Imaging 6, 236-237.

Guinier A. 1963. "X-ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies", W.H. Freeman, San Fransisco.

Hanstead P.D. 1981. "Simplified digital synthesis of ultrasonic Images", Proc. R. Soc. (Lond.) A374, 491-502.

Harrington R.F. 1968. "Field Computations by Moment Methods", Macmillan, New York.

Harris F.J. 1978. "On the use of windows for harmonic analysis with the discrete Fourier Transform", Proc. IEEE 66, 51-80.

Hawkes P.W. 1983a. "Computer-aided calculation of the aberration coefficient of microwave cavity lenses. Part I, primary (second-order) aberrations", Optik 63, 129-156.

Hawkes P.W. 1983b. "Computer-aided calculation of the aberration coefficient of microwave cavity lenses. Part II, secondary (third-order) aberrations", Optik 65, 227-251.

Heckman R. and Seidel K.J. 1983. "The sonographic appearance and contrast enhancement of puncture needles", J. Clin. Ultrasound 11, 265-268.

Heffernan P.B. and Bates R.H.T. 1982. "Image reconstruction from projections VI: Comparision of interpolation methods", Optik 50, 129-142.

Hefner L.V., Parks J.A. and Goldstein A. 1980. "Transducer beam pattern test object", J. Clin. Ultrasound 8, 5-10.

Herstein I.N. 1975. "Topics in Algebra", J. Wiley and Sons, New York, 2nd. Edition.

Hildebrand B.P. 1980. "An analysis of pulsed ultrasonic arrays", Acoustical Imaging 8, 165-184.

Hundt E.E. and Trautenberg E.A. 1980. "Digital processing of ultrasonic data by deconvolution", IEEE Trans. SU-27, 249-252.

Hunt B.R., Fright W.R. and Bates R.H.T. 1983. "Analysis of the shift-and-add method for imaging through turbulent media", J. Opt. Soc. Am. 73, 456-465.

Ishimaru A. 1978. "Wave Propagation and Scattering in Random Media", Academic Press, New York, 2 volumes.

Iwata K. and Nagata R. 1975. "Calculation of refractive index distribution from interferograms using the Born and Rytov's approximation", Jap. J. Phys. (suppl. 14-1) 14, 379-383.

James G.L. 1976. "Geometrical Theory of Diffraction for Electromagnetic Waves", IEE Electromagnetic Waves Series 1, G. Millington, E.D.R Shearman and J.R. Wait Ed., Peter Peregrinus Ltd., Stevenage.

Jellins J., Kossoff G., Reeve T.S. and Barraclough B.H. 1975. "Ultrasonic grey scale visualisation of breast disease", Ultrasound in Med. & Biol. 1, 393-404.

Johnson S.A., Greenleaf J.F., Samayoa W.F., Duck F.A. and Sjostrand J.D. 1975. "Reconstruction of three dimensional velocity fields and other parameters by acoustic ray tracing", IEEE Ultrasonics Symposium Proceedings, IEEE catalog no. 75 CHO 994-4SU, 46-51.

Johnson R.K. and Devaney A.J. 1982. "Transducer effects in acoustic scattering measurements", Appl. Phys. Lett. 41, 622-624.

Jones D.S. 1964. "Theory of Electromagnetism", Pergamon Press, Oxford.

Karrer H.E., Dias J.F., Larsen J.D. and Pering R.D. 1982. "A phased array acoustic imaging system for medical use", Acoustical Imaging 10, 47-63.

Kaveh M., Soumekh M. and Mueller R.K. 1982. "A comparision of Born and Rytov approximations in acoustic tomography", Acoustical Imaging 11, 325-335.

Kaveh M., Soumekh M. and Greenleaf J.F. 1984. "Signal processing for diffraction tomography", IEEE Trans. SU-31, 230-239.

Kerns D.M. 1975. J. Acoust. Soc. Am. 57, 497.

Keller J.B., Lewis R.M. and Seckler B.D. 1956. "Asymptotic solution of some diffraction problems", Comm. Pure Appl. Maths. 9, 207-265.

Keller J.B. 1969. "Accuracy and validity of the Born and Rytov approximations", J. Opt. Soc. Am. 59, 1003-1004.

Kleinman R.E. 1965. "The Rayleigh region", Proc. IEEE 53, 848-856.

Klepper J.R., Brandenburger G.M., Mimbs J.W., Sobel B.E. and Miller J.G. 1981. "Application of phase-insensitive detection and frequency dependent measurements to computed ultrasonic attenuation tomography", IEEE Trans. BME-28, 186-201.

Knox K.J. and Thompson B.J. 1974. "Recovery of images from atmospherically degraded short-exposure photographs", Astrophys. J. 193, L45-L48.

Kobayashi T. 1979. "Diagnostic ultrasound in breast cancer: Analysis of retrotumorous echo patterns correlated with sonic attenuation by cancerous connective tissue", J. Clin. Ultrasound 7, 471-479.

Kossoff G., Robinson D.E. and Garrett W.J. 1968. "Ultrasonic two-dimensional visualisation for medical diagnosis", J. Acoust. Soc. Am. 44, 1310-1318.

Kossoff G. 1973. "Ultrasonic research in medicine in Australia: A review", Ultrasonics International 1973 (Conference Proceedings), I.P.C. Science and Technology Press, Guildford U.K., 199-205.

Kossoff G., Garrett W.J., Carpenter D.A., Jellins J. and Dadd M.J. 1976. "Principles and classification of soft tissues by grey scale echography", Ultrasound in Med. & Biol. 2, 89-105.

Kouyoumjian R.G. 1965. "Asymptotic high-frequency methods", Proc. IEEE 53, 864-876.

Kuc R. 1984. "Feasibility of attenuation images from reflected ultrasound signals", 9th International Symposium on Ultrasonic Imaging and Tissue Characterisation 1984 (Abstract), in Ultrasonic Imaging 6, 213.

Kunz K.S. and Lee K.M. 1978a. "A three-dimensional finite-difference solution of the external response of an aircraft to a complex transient EM enviroment: Part I - the method and its implementation", IEEE Trans. EMC-20, 328-333.

Kunz K.S. and Lee K.M. 1978b. "A three-dimensional finite-difference solution of the external response of an aircraft to a complex transient EM enviroment: Part II - comparision of predictions and measurements", IEEE Trans. EMC-20, 333-341.

Labeyrie A. 1970. "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images", Astron. Astrophys. 6, 85-87.

Larsen M.D. 1969. "Introduction to Modern Algebraic Concepts", Addison-Wesley, Reading, Massachusetts.

Lee S.W. 1978. "Uniform Asymptotic Theory of Electromagnetic Edge Diffraction: A Review", in Electromagnetic Scattering, P.L.E. Uslenghi Ed., Academic Press, New York.

Leeman S., Ferrari L., Jones J.P. and Fink M. 1984. "Perspectives on attenuation estimation from pulse-echo signals", IEEE Trans. SU-31, 351-361.

Lewin L. 1970a. "On the inadequacy of discrete mode-matching techniques in some waveguide discontinuity problems", IEEE Trans. MTT-18, 364-372.

Lewin L. 1970b. "On the restricted validity of point-matching techniques", IEEE Trans. MTT-18, 1041-1047.

Lewitt R.M. 1977. "Contributions to image reconstruction", Ph.D. Thesis, University of Canterbury, New Zealand.

Lewitt R.M. and Bates R.H.T. 1978. "Image reconstruction from projections I: General theoretical considerations", Optik 50, 19-33.

Lewitt R.M., Bates R.H.T. and Peters T.M. 1978. "Image reconstruction from projections II: Modified back-projection method", Optik 50, 85-109.

Ligtwoet G., Rijsterborgh H., Kappen L. and Bom N. 1978. "Real time ultrasonic imaging with a hand held scanner", Ultrasound in Med. & Biol. 4, 91-98.

Linzer M. and Parks S.I. 1979. "The sonochromascope", Computer Aided Tomography and Ultrasonics in Medicine, J. Raviv, J.F. Greenleaf and G.T. Herman Ed., North-Holland, Amsterdam, 73-75.

Lohmann A.W., Weigelt G. and Wirnitzer B. 1983. "Speckle masking in astronomy: triple correlation theory and applications", Appl. Opt. 22, 4028-4037.

Longley L.A. and O'Brien W.D. (Jr.) 1982. "Ultrasonic heating distribution in lossy cylinders and spheres", IEEE Trans. SU-29, 66-79.

Lu Z.Q., Kaveh M. and Mueller R.K. 1984. "Diffraction tomography using beam waves: Z-average reconstruction", Ultrasonic Imaging 6, 95-102.

Lynds C.R., Worden S.P. and Harvey J.W. 1976. "Digital image reconstruction applied to Alpha Orionis", Astrophys. J. 207, 174-180.

Macovski A. and Norton S.J. 1975. "High resolution B-scan systems using a circular array", Acoustical Holography 6, 121-143.

Macovski A. 1976. "Theory on imaging with arrays", Acoustic Imaging G. Wade Ed., Plenum Press, New York, 111-126.

Macovski A. 1983. "Medical Imaging Systems", Prentice-Hall, Englewood Cliffs, N.J.

Maginness M.G., Plummer J.D., Beaver W.L. and Meindl J.D. 1976. "State-of-the-art in two-dimensional ultrasonic transducer array technology", Med. Phys. 3, 312-318.

Maginness M.G. 1979. "Methods and terminology for diagnostic ultrasound imaging systems", Proc. IEEE 67, 641-653.

McDonnell M.J. and Bates R.H.T. 1975. "Restoring parts of scenes from blurred photographs", Opt. Commun. 13, 347-349.

McKinnon G.C. 1980. "Contributions to Imaging", Ph.D. Thesis, University of Canterbury, New Zealand.

McKinnon G.C. and Bates R.H.T. 1980. "A limitation on ultrasonic transmission tomography", Ultrasonic Imaging 2, 48-54.

McKinnon G.C., Schomberg H., Biel W., Proksa R. and Tschendel O. 1984. "Improvements to ultrasonic transmission imaging of the breast", 9th International Symposium on Ultrasonic Imaging and Tissue Characterisation 1984 (Abstract), in Ultrasonic Imaging 6, 234.

Melton H.E. (Jr.) and Magnin P.A. 1984. "A-mode speckle reduction with compound frequencies and compound bandwidths", Ultrasonic Imaging 6, 159-173.

Middleton D.M. 1960. "An Introduction to Statistical Communication Theory", McGraw-Hill, New York.

Miller E.B., Smith S.W. and Thurstone F.L. 1974. "A study of near field ultrasonic beam patterns from a pulsed linear array", Acoustical Holography 5, 261-282.

Milner M.O. 1979. "Error correction in images and imaging instruments", Ph.D. Thesis, University of Canterbury, New Zealand.

Minard R.A., Robinson B.S. and Bates R.H.T. 1985. "Full-wave computed tomography. Part 3: Coherent shift-and-add imaging", Proc. IEE 132, 50-58.

Miyashita T., Makayama J. and Okura H. 1980. "Acoustical imaging by means of multi-frequency hologram matrix", Acoustical Imaging 9, 23-37.

Mole R.H. 1978. "The sensitivity of the human breast to cancer induction by ionising radiation", Brit. J. Radiology 51, 401-405.

Morgan C.L., Trought W.S., Clark W.M., von Ramm O.T. and Thurstone F.L. 1978. "Principles and applications of a dynamically focused phased array real time ultrasound system", J. Clin. Ultrasound 6, 385-391.

Morita N. 1978. "Surface Integral Representations for Electromagnetic Scattering from Dielectric Cylinders", IEEE Trans. AP-26, 261-266.

Morse P.M. and Feshbach H. 1953. "Methods of Mathematical Physics", McGraw-Hill, New York, 2 volumes.

Morse P.M. and Ingard K.U. 1968. "Theoretical Acoustics", McGraw-Hill, New York.

Mueller R.K., Kaveh M. and Wade G. 1979. "Reconstructive tomography and applications to ultrasonics", Proc. IEEE 67, 567-586.

Mueller R.K. 1980. "Diffraction tomography I: The wave equation", Ultrasonic Imaging 2, 213-222.

Mur G. 1981a. "The modelling of singularities in the finite-difference approximations of the time-domain electromagnetic-field equations", IEEE Trans. MTT-29, 1073-1077.

Mur G. 1981b. "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic-field equations", IEEE Trans. EMC-23, 377-382.

Nahamoo D. and Kak A.C. 1982. "Ultrasonic Diffraction Imaging", School of Electrical Engineering, Purdue University, Tech. Rep. TR-EE-82-20.

Nahamoo D., Pan S.X. and Kak A.C. 1984. "Synthetic aperture diffraction tomography and its interpolation-free computer implementation", IEEE Trans. SU-31, 218-229.

Nash J.C. 1979. "Compact numerical methods for computers: Linear algebra and function minimisation", Adam Hilger Ltd., Bristol.

Nhunduru K.E. 1984. "Software and Hardware for Ultrasonic Imaging", 3rd Pro. Project Report, Electrical and Electronic Engineering Department, University of Canterbury, New Zealand.

Nitadori K. 1975. "Synthetic aperture approach to multi-beam scanning acoustical imaging", Acoustical Holography 6, 507-523.

Nitadori K., Mano K. and Kamata H. 1980. "An experimental underwater acoustic imaging using multi-beam scanning", Acoustical Imaging 8, 249-266.

Norton S.J. and Linzer M. 1981. "Ultrasonic reflectivity imaging in three dimensions: exact inverse scattering solutions for plane, cylindrical, and spherical apertures", IEEE Trans. BME-28, 202-220.

Norton S.J. 1983. "Generation of separate density and compressibility images in tissue", Ultrasonic Imaging 5, 240-252.

Ophir J. and Maklad J.F. 1979. "Digital scan converters in diagnostic ultrasound imaging", Proc. IEEE 67, 654-664.

Oristaglio M.L. 1983. "A geometric approach to the filtered backpropagation algorithm", Ultrasonic Imaging 5, 30-37.

Pan S.X. and Kak A.C. 1983. "A computational study of reconstruction algorithms for diffraction tomography: Interpolation versus filtered-backpropagation", IEEE Trans. ASSP-31, 1262-1276.

Papadakis E.P. 1975. "Ultrasonic diffraction from single apertures with application to pulse measurements and crystal physics" in Physical Acoustics, W.P. Mason and R.N. Thurston Eds., Academic Press, New York, 151-211.

Patterson M.S. and Foster F.S. 1983. "The improvement and quantitative assessment of B-mode images produced by an annular array/cone hybrid", Ultrasonic Imaging 5, 195-213.

Pavelle R., Rothstein M. and Fitch J. 1981. "Computer Algebra", Scientific American, 102-113.

Peters T.M. 1973. "Image reconstruction from projections", Ph.D. Thesis, University of Canterbury, New Zealand.

Petersen D.K. and Kino G.S. 1984. "Real-time digital image reconstruction: a description of imaging hardware and an analysis of quantization errors", IEEE Trans. SU-31, 337-351.

Plummer J.D., Swartz R.G., Maginness M.G., Beadouin J.R. and Meindl J.D. 1978. "Two-dimensional transmit/receive ceramic piezoelectric arrays: Construction and performance", IEEE Trans. SU-25, 273-280.

Pogorzelski W.A. 1966. "Integral Equations and their Applications", Pergammon Press, Oxford.

Pohlhammer J. and O'Brien W.D. (Jr.) 1981. "Dependence of the ultrasonic scatter coefficient on collagen concentration in mammalian tissues", J. Acoust. Soc. Am. 69, 283-285.

Prabhu K.M.N., Reddy V.U. and Agrawal J.P. 1977. "Performance comparison of data windows", Electron. Lett. 13, 600-601.

Reintjes J.F. and Coate G.T. 1952. "Principles of Radar", McGraw-Hill, New York, 3rd. Edition.

Rihaczek A.W. 1969. "Principles of High Resolution Radar", McGraw-Hill, New York.

Robinson D.E., Jellins J. and Kossoff G. 1970. "The C.A.L. four channel solid state echo-encephaloscope", Ultrasonics 8, 93-96.

Robinson D.E., Lees S. and Bess L. 1974. "Near field transient radiation patterns for circular pistons", IEEE Trans. ASSP-22, 395-403.

Robinson B.S. and Bates R.H.T. 1980. "Wideband ultrasonic diffraction measurements", Australian Physical Sciences and Medicine 3, 233-238.

Robinson D.E. and Knight P.C. 1981. "Computer reconstruction techniques in compound scan pulse-echo imaging", Ultrasonic Imaging 3, 217-234.

Robinson B.S. 1982. "Speckle Processing for Ultrasonic Imaging", Ph.D. Thesis, University of Canterbury, New Zealand.

Robinson D.E., Chen C.F. and Wilson L.S. 1983. "Image matching for pulse echo measurement of ultrasonic velocity", Image and Vision Computing 1, 145-151.

Robinson D.E. 1984. "Digital reconstruction and display of compound scan ultrasound images", IEEE Trans. SU-31, 396-406.

Robinson D.E., Wilson L.S. and Bianchi T. 1984. "In vivo characterisation of beams for ultrasonic attenuaton measurements", 9th International Symposium on Ultrasonic Imaging and Tissue Characterisation 1984 (Abstract), in Ultrasonic Imaging 6, 217.

Robinson D.E. and Wing M. 1984. "Lateral deconvolution of ultrasonic beams", Ultrasonic Imaging 6, 1-12.

Robinson B.S. and Greenleaf J.F. 1985. "Computerized ultrasound tomography", Ch. VIII in "3-D Imaging Methods in Medicine and Biology", R.A. Robb Ed., CRC Press, Boca Raton, Florida.

Sapriel J. 1979. "Acousto-optics", J. Wiley and Sons, Chichester.

Sato T. and Wadaka S. 1979. "Ultrasonic imaging system using spatial and temporal M-sequences", Ultrasonic Imaging 1, 245-254.

Schmitt R., Meyer C.R., Carson P.L., Chenevert T.L. and Bland P.H. 1984. "Error reduction in through transmission tomography using large receiving arrays with phase-insensitive signal processing", IEEE Trans. SU-31, 251-258.

Schomberg H. 1978. "An improved approach to reconstructive ultrasound tomography", J. Phys. D. (Appl. Phys.) 11, L181-L185.

Schomberg H. 1982. "Nonlinear image reconstruction from ultrasonic time-of-flight projections", Acoustical Imaging 10, 381-396.

Schomberg H., Vollman W. and Mahnke G. 1983. "Lateral inverse filtering of ultrasonic B-scan images", Ultrasonic Imaging 5, 38-54.

Seki H., Granato A. and Truell R. 1956. J. Acoust. Soc. Am. 28, 230.

Selbie R.D., Hutchison J.M.S. and Mallard J.R. 1980. "The Aberdeen phased array: A real time ultrasonic scanner with dynamic focus", Med. and Biol. Eng. and Comp. 18, 335-343.

Seckler B.D. and Keller J.B. 1959. "Geometrical theory of diffraction in inhomogeneous media", J. Acoust. Soc. Amer. 31, 192-206.

Silver S. (Ed.) 1965. "Microwave Antenna Theory and Design", Dover, New York.

Silvester P. 1969. "A general high-order finite-element waveguide analysis program", IEEE Trans. MTT-17, 204-210.

Slaney M. and Kak A.C. 1983. "Diffraction Tomography", Proc. SPIE Conf. Inverse Optics 413, 2-19.

Smith S.W., von Ramm O.T., Haran M.E. and Thurstone F.L. 1979. "Angular response of piezoelectric elements in phased array ultrasound scanners", IEEE Trans. SU-26, 185-190.

Smith S.W., Trahey G.E. and von Ramm O.T. 1984. "Refraction correction for ultrasonic imaging through tissue layers", 9th International Symposium on Ultrasonic Imaging and Tissue Characterisation 1984 (Abstract), in Ultrasonic Imaging 6, 210-211.

Sokollu A. 1971. "Irreversible effects of high frequency ultrasound on animal tissue and related threshold intensities", Acoustical Holography 3, 3-18.

Stark H. and Paul I. 1981. "An investigation of computerized tomography by direct Fourier inversion and optimum interpolation", IEEE Trans. BME-28, 496-505.

Steinberg B.D. 1976. "Principles of Aperture and Array System Design", J. Wiley and Sons, New York.

Steinberg B.D. and Luthra A.K. 1982. "Sidelobe reduction of the ring array for use in circularly symmetric imaging systems", Acoustical Imaging 11, 479-490.

Stern R.A., Hill B.C., Gaudiani V.A., Green S.E. and Ingels N.B. (Jr.) 1983. "A biologically compatible implantable ultrasonic marker", Ultrasound in Med. & Biol. 9, 191-200.

Stockham T.G. (Jr.), Cannon T.M. and Ingebretsen R.B. 1975. "Blind deconvolution through digital signal processing", Proc. IEEE 63, 678-692.

Swartz R.G. and Plummer J.D. 1980. "On the generation of high-frequency acoustic energy with polyvinylidene fluoride", IEEE Trans. SU-27, 295-302.

Szuszczewicz E.P., Rodriguez P., Singh M and Mango S. 1983. "Ionospheric irregularities and their potential impact on synthetic aperture radars", Radio Sci. 18, 765-774.

Takeuchi Y. 1979 "An investigation of a spread energy method for medical ultrasound systems (Parts 1 & 2)", Ultrasonics 17, 175-182, 219-225.

Thurstone F.L. and von Ramm O.T. 1974. "A new ultrasound imaging technique employing two-dimensional electronic beam steering", Acoustical Holography 5, 249-259.

Ufimstev P. Ya. 1962. "The method of fringe waves in the physical theory of diffraction", Sovyetskoye radio, Moscow, 1-143.

Available in translation from the U.S Air Force Foreign Technology Division, Wright-Patterson, AFB, Ohio, USA.

Ursell F. 1966. "On the rigorous foundation of short-wave asymptotics", Proc. Camb. Phil. Soc. 62, 227-244.

Vakine R. and Lorenz W.J. 1984. "Lateral filtering of medical ultrasonic B-scans before image generation", Ultrasonic Imaging 6, 152-158.

Vilkomerson D. and Hurley B. 1975 "Progress in annular-array imaging", Acoustical Holography 6, 145-164.

von Ramm O.T. and Smith S.W. 1978. "A multiple frequency array for improved diagnostic imaging", IEEE Trans. SU-25, 340-345.

Wells P.N.T. 1977. "Biomedical Ultrasonics", Academic Press, London.

Wells P.N.T. and Halliwell M. 1981. "Speckle in ultrasonic imaging", Ultrasonics 19, 225-229.

Wesolowski K. 1983. "Computer generation of a slowly pseudorandom process", IEE Proc. F 130, 314-316.

Weyns A. 1980. "Radiation field calculations of pulsed ultrasonic transducers (Parts 1 & 2)", Ultrasonics 18, 183-188, 219-223.

Wolf E. 1969. "Three-dimensional structure determination of semi transparent objects from holographic data", Opt. Commun. 1, 153-156.

Yamamoto T. and Aoki Y. 1984. "Holographic B-scope imaging using wide-band chirped ultrasound", IEEE Trans. SU-31, 362-366.

Zabreyko P.P., Koshelev A.I., Krasnosel'skii M.A., Mikhlin S.G., Rakovshchik L.S. and Stet'senko V. Ya. 1975. "Integral Equations - A Reference Text", Noordhoff, Leyden.